GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:32:00 ; Search time 10277.7 Seconds
(without alignments)
4505.438 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1118133 seqs, 2558875100 residues

Total number of hits satisfying chosen parameters: 2236266

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_om:*
4: gb_ov:*
5: gb_ph:*
6: gb_pl1:*
7: gb_pl2:*
8: gb_pr1:*
9: gb_pr2:*
10: gb_pr3:*
11: gb_ro:*
12: gb_sy:*
13: gb_un:*
14: em_fun:*
15: em_hum1:*
16: em_hum2:*
17: em_in:*
18: em_om:*
19: em_or:*
20: em_ov:*
21: em_pat:*
22: em_ph:*
23: em_pl:*
24: em_ro:*
25: em_sts:*
26: em_sy:*
27: em_un:*
28: em_vi:*
29: gb_ba3:*
30: gb_in1:*
31: gb_in2:*
32: gb_in3:*
33: gb_pl3:*
34: gb_pr4:*
35: em_ba1:*
36: em_ba2:*
37: em_htg1:*
38: em_htg2:*
39: em_htg3:*
40: em_htg4:*
41: em_htg5:*
42: em_htg6:*
43: em_htg7:*
44: em_htg8:*
45: em_htg9:*
46: em_htg10:*
47: em_hum3:*
48: em_hum4:*
49: em_hum5:*
50: em_hum6:*
51: gb_pr5:*
52: gb_pr6:*
53: gb_pr7:*
54: gb_htg1:*
55: gb_htg2:*
56: gb_htg3:*
57: gb_htg4:*
58: gb_htg5:*
59: gb_htg6:*
60: gb_htg7:*
61: gb_htg8:*
62: gb_htg9:*
63: gb_htg10:*
64: gb_htg11:*
65: gb_htg12:*
66: gb_htg13:*
67: gb_htg14:*
68: gb_htg15:*
69: gb_htg16:*
70: gb_htg17:*
71: gb_htg18:*
72: gb_htg19:*
73: gb_htg20:*
74: gb_htg21:*
75: gb_htg22:*
76: gb_htg23:*
77: gb_sts1:*
78: gb_sts2:*
79: gb_vi1:*
80: gb_vi2:*
81: gb_pat1:*
82: gb_pat2:*
83: em_htg0:*
84: gb_htg24:*
85: gb_pr8:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		9046.8	100.0	9108	10	AF123653	AF123653 Homo sapi
2	c	4926.8	54.5	172162	64	AC025853	AC025853 Homo sapi
3		4199.8	46.4	5492	10	AF123659	AF123659 Homo sapi
4		807.2	8.9	1515	10	AF123656	AF123656 Homo sapi
5		807.2	8.9	1614	10	AF123655	AF123655 Homo sapi
6		807.2	8.9	1692	10	AF123657	AF123657 Homo sapi
7		807.2	8.9	1722	10	AF123658	AF123658 Homo sapi
8		346	3.8	633	10	AF123654	AF123654 Homo sapi
9	c	229.2	2.5	163683	64	AC025412	AC025412 Homo sapi
10		226.2	2.5	152511	60	AC019269	AC019269 Homo sapi
11		225.6	2.5	179538	63	AC024740	AC024740 Homo sapi
12	c	224.6	2.5	159440	34	AL355392	AL355392 Human DNA
13		224.2	2.5	177560	59	AC016695	AC016695 Homo sapi
14		224	2.5	227949	65	AC026803	AC026803 Homo sapi
15		223.6	2.5	160659	66	AC036235	AC036235 Homo sapi
16		223.6	2.5	180516	84	CNS01DSH	AL121784 Homo sapi
17		223.6	2.5	190128	54	AC007374	AC007374 Homo sapi
18		223.6	2.5	215945	74	AL355337	AL355337 Homo sapi
19	c	223.4	2.5	89818	8	AC002126	AC002126 Homo sapi
20	c	223.2	2.5	141292	72	AF235106	AF235106 Homo sapi
21	c	223	2.5	156345	62	AC023015	AC023015 Homo sapi

	22	223	2.5	203088	73	AL162505	AL162505 Homo sapi
c	23	222.6	2.5	189623	60	AC018640	AC018640 Homo sapi
c	24	222.6	2.5	189663	55	AC009244	AC009244 Homo sapi
c	25	222.4	2.5	170664	74	AL358013	AL358013 Homo sapi
.	26	222.2	2.5	191162	9	AC008865	AC008865 Homo sapi
c	27	222	2.5	162176	72	AL137780	AL137780 Homo sapi
c	28	222	2.5	169236	69	AC073167	AC073167 Homo sapi
c	29	222	2.5	251664	55	AC008908	AC008908 Homo sapi
c	30	221.8	2.5	110000	55	AC008749_0	AC008749 Homo sapi
.	31	221.8	2.5	154621	75	AL391836	AL391836 Homo sapi
c	32	221.6	2.4	184316	63	AC024022	AC024022 Homo sapi
	33	221.2	2.4	188740	63	AC024045	AC024045 Homo sapi
	34	221	2.4	175839	8	AC004491	AC004491 Homo sapi
	35	220.6	2.4	170866	64	AC026203	AC026203 Homo sapi
c	36	220.6	2.4	189268	63	AC024021	AC024021 Homo sapi
c	37	220.4	2.4	151217	64	AC026164	AC026164 Homo sapi
c	38	220.4	2.4	160095	59	AC018458	AC018458 Homo sapi
	39	220	2.4	158859	66	AC040895	AC040895 Homo sapi
	40	220	2.4	191979	54	AC005867	AC005867 Homo sapi
	41	219.8	2.4	208202	57	AC013446	AC013446 Homo sapi
	42	219.6	2.4	180638	72	AL157709	AL157709 Homo sapi
c	43	219.6	2.4	278310	69	AC069592	AC069592 Homo sapi
c	44	219.2	2.4	144293	62	AC022408	AC022408 Homo sapi
	45	219.2	2.4	156895	57	AC012652	AC012652 Homo sapi

ALIGNMENTS

```
RESULT   1
AF123653
LOCUS       AF123653     9108 bp    DNA             PRI       07-APR-1999
DEFINITION  Homo sapiens FEZ1 (FEZ1) gene, complete cds.
ACCESSION   AF123653
VERSION     AF123653.1  GI:4572463
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 9108)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 9108)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .9108
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
     mRNA            join(<112. .456,1707. .2510,4912. .>5553)
                     /gene="FEZ1"
                     /product="FEZ1"
     gene            <112. .>5553
                     /gene="FEZ1"
     CDS             join(112. .456,1707. .2510,4912. .5553)
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23834.1"
                     /db_xref="GI:4572464"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQRKKNEAELLREKVNLLEQELQELRAQAALARDMGPPTFPEDVPALQRELERLR
                     AELREERQGHDQMSSGFQHERLVWKEEKEKVIQYQKQLQQSYVAMYQRNQRLEKALQQ
                     LARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT     2021 a   2628 c   2504 g   1952 t      3 others
ORIGIN
```

Query Match 100.0%; Score 9046.8; DB 10; Length 9108;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9048; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 gcctttccaagaccctgcccggccctgccccatcctcagccccgagtcaccatgggcagc 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 GCCTTTCCAAGACCCTGCCCGGCCCTGCCCCATCCTCAGCCCCGAGTCACCATGGGCAGC 120

Qy      61 gtcagtagcctcatctccggccacagcttccacagcaagcactgccgggcttcgcagtac 120
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 GTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCTTCGCAGTAC 180

Qy     121 aagctgcgcaagtcctcccacctcaagaagctcaaccggtattccgacgggctgctgagg 180
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGCTGCGCAAGTCCTCCCACCTCAAGAAGCTCAACCGGTATTCCGACGGGCTGCTGAGG 240

Qy     181 tttggcttctcccaggactccggtcacggcaagtccagctccaaaatgggcaagagcgaa 240
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 TTTGGCTTCTCCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAAATGGGCAAGAGCGAA 300

Qy     241 gacttcttctacatcaaggtcagccagaaagcccggggctcccatcacccagattacacg 300
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     301 GACTTCTTCTACATCAAGGTCAGCCAGAAAGCCCGGGGCTCCCATCACCCAGATTACACG 360

Qy     301 gcactgtccagcggggatttaggggccaggctggggtggactttgacccgtccacaccc 360
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     361 GCACTGTCCAGCGGGGATTTAGGGGCCAGGCTGGGGTGGACTTTGACCCGTCCACACCC 420

Qy     361 cccaagctcatgcccttctccaatcagctagaaatggtaagcgggggtcgctggcaaggg 420
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     421 CCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGGTAAGCGGGGGTCGCTGGCAAGGG 480

Qy     421 taagtgggttggaaacgcaggagaaagcaaaatgggggtggagagcctgggggttcaggg 480
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     481 TAAGTGGGTTGGAAACGCAGGAGAAAGCAAAATGGGGGTGGAGAGCCTGGGGGTTCAGGG 540

Qy     481 ggagtggtgacctgagcattcagactcctcaaaaccagagcggcaggggtgccggcggaa 540
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     541 GGAGTGGTGACCTGAGCATTCAGACTCCTCAAAACCAGAGCGGCAGGGGTGCCGGCGGAA 600

Qy     541 gcctgtggccacaccgcagagatcaaacgtttcacaaaggaattagagcatcgctcagtc 600
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     601 GCCTGTGGCCACACCGCAGAGATCAAACGTTTCACAAAGGAATTAGAGCATCGCTCAGTC 660

Qy     601 cccctgaagcagaagtcttgggtcaggccataagcaaagagcacaggggatatgtgagct 660
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     661 CCCCTGAAGCAGAAGTCTTGGGTCAGGCCATAAGCAAAGAGCACAGGGGATATGTGAGCT 720

Qy     661 tttggagtcccactgaaatgtagctggattgtcaacgtaggatccaggcgtttgccaagc 720
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     721 TTTGGAGTCCCACTGAAATGTAGCTGGATTGTCAACGTAGGATCCAGGCGTTTGCCAAGC 780

Qy     721 ctcgggaaggagagggagccctgttctcatctggaagcacagatgaagaggatgcaggcc 780
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     781 CTCGGGAAGGAGAGGGAGCCCTGTTCTCATCTGGAAGCACAGATGAAGAGGATGCAGGCC 840

Qy     781 gggagttaaccgcttctctccccgggagactcgtgggggtgggtgcggtcttctcatttg 840
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     841 GGGAGTTAACCGCTTCTCTCCCCGGGAGACTCGTGGGGGTGGGTGCGGTCTTCTCATTTG 900

Qy     841 ctgccctggtgtgcattagctccttgttcaagctgcgcctgggggcatctttgaatacag 900
```

```
Db    901 CTGCCCTGGTGTGCATTAGCTCCTTGTTCAAGCTGCGCCTGGGGGCATCTTTGAATACAG 960

Qy    901 gctggagttttgtcatccatttaccagagactagggcaaaggaggcccaggcactgagaa 960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    961 GCTGGAGTTTTGTCATCCATTTACCAGAGACTAGGGCAAAGGAGGCCCAGGCACTGAGAA 1020

Qy    961 atccagccctcacaccagctcaagccctcgtgcgtcccacgagtggacactgaaatcaat 1020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1021 ATCCAGCCCTCACACCAGCTCAAGCCCTCGTGCGTCCCACGAGTGGACACTGAAATCAAT 1080

Qy   1021 tttcctattcagtcctctgccccttgccctggggaaatgaatccccggctttgatttact 1080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1081 TTTCCTATTCAGTCCTCTGCCCCTTGCCCTGGGGAAATGAATCCCCGGCTTTGATTTACT 1140

Qy   1081 aggaaagagcctcttatgtttgcatagagcattcagcttttcaaattaaggggcttgtaa 1140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1141 AGGAAAGAGCCTCTTATGTTTGCATAGAGCATTCAGCTTTTCAAATTAAGGGGCTTGTAA 1200

Qy   1141 actgtgaagcactctaccagggaaaattacagttttaaaaaaggatcgtgatttggagtg 1200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1201 ACTGTGAAGCACTCTACCAGGGAAAATTACAGTTTTAAAAAAGGATCGTGATTTGGAGTG 1260

Qy   1201 agcctcccaaccctgtaaggaggccaggtccgtgtccttgctccaggcttaatggaagag 1260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1261 AGCCTCCCAACCCTGTAAGGAGGCCAGGTCCGTGTCCTTGCTCCAGGCTTAATGGAAGAG 1320

Qy   1261 gcagtgaacaggaagaagggatggacctaaagagggacagcaagctcggccagcctgatg 1320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1321 GCAGTGAACAGGAAGAAGGGATGGACCTAAAGAGGGACAGCAAGCTCGGCCAGCCTGATG 1380

Qy   1321 ccctaacttgccccacacagagacctagagcaggagcctcaagatggtatttatcacctc 1380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1381 CCCTAACTTGCCCCACACAGAGACCTAGAGCAGGAGCCTCAAGATGGTATTTATCACCTC 1440

Qy   1381 gggagggctggggcaagctggtggcaggttgctatttcatagaacaaagtgcccaagtcg 1440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1441 GGGAGGGCTGGGGCAAGCTGGTGGCAGGTTGCTATTTCATAGAACAAAGTGCCCAAGTCG 1500

Qy   1441 ccattagggtttttccctcctaagagagatgacattcagctgcttcaaagcaacaggcaa 1500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1501 CCATTAGGGTTTTTCCCTCCTAAGAGAGATGACATTCAGCTGCTTCAAAGCAACAGGCAA 1560

Qy   1501 ggtctgctgagacaattgaccaagaggggtgctgcgtgcgctcagagagcccagactggc 1560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1561 GGTCTGCTGAGACAATTGACCAAGAGGGGTGCTGCGTGCGCTCAGAGAGCCCAGACTGGC 1620

Qy   1561 tcaaggtcggcacgcgtgcctggggagggagggtgcaatgcgcgcgcaggggaggcatga 1620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1621 TCAAGGTCGGCACGCGTGCCTGGGGAGGGAGGGTGCAATGCGCGCGCAGGGGAGGCATGA 1680

Qy   1621 gtcaccgcggtccttttcctctacagggctccgagaagggtgcagtgaggcccacagcct 1680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1681 GTCACCGCGGTCCTTTTCCTCTACAGGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCT 1740

Qy   1681 tcaagcctgtgctgccacggtcaggagccatcctgcactcctccccggagagtgccagcc 1740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1741 TCAAGCCTGTGCTGCCACGGTCAGGAGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCC 1800

Qy   1741 accagctgcaccccgcccctccagacaagcccaaggagcaggagctgaagcctggcctgt 1800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1801 ACCAGCTGCACCCCGCCCCTCCAGACAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGT 1860

Qy   1801 gctctggggcgctgtcagactccggccggaactccatgtccagcctgcccacacacagca 1860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1861 GCTCTGGGGCGCTGTCAGACTCCGGCCGGAACTCCATGTCCAGCCTGCCCACACACAGCA 1920

Qy   1861 ccagcagcagctaccagctggacccgctggtcacacccgtgggacccacaagccgttttg 1920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1921 CCAGCAGCAGCTACCAGCTGGACCCGCTGGTCACACCCGTGGGACCCACAAGCCGTTTTG 1980

Qy   1921 ggggctccgcccacaacatcacccagggcatcgtcctccaggacagcaacatgatgagcc 1980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1981 GGGGCTCCGCCCACAACATCACCCAGGGCATCGTCCTCCAGGACAGCAACATGATGAGCC 2040

Qy   1981 tgaaggctctgtccttctccgacggaggtagcaagctgggccactcgaacaaggcagaca 2040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2041 TGAAGGCTCTGTCCTTCTCCGACGGAGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACA 2100

Qy   2041 agggcccctcgtgtgtccgctcccccatctccacggacgagtgcagcatccaggagctgg 2100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2101 AGGGCCCCTCGTGTGTCCGCTCCCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGG 2160

Qy   2101 aacagaagctgttggagagggagggcgccctccagaagctgcagcgcagctttgaggaga 2160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2161 AACAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGA 2220

Qy   2161 aggagcttgcctccagcctggcctacgaggagcggccgcggcgctgcagggacgagctgg 2220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2221 AGGAGCTTGCCTCCAGCCTGGCCTACGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGG 2280

Qy   2221 agggcccggagcccaaaggcggcaacaagctcaagcaggcctcgcagaagagccagcgcg 2280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2281 AGGGCCCGGAGCCCAAAGGCGGCAACAAGCTCAAGCAGGCCTCGCAGAAGAGCCAGCGCG 2340

Qy   2281 cgcagcaggtcctgcacctgcaggtactgcagcttcagcaggagaagcggcagctccggc 2340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2341 CGCAGCAGGTCCTGCACCTGCAGGTACTGCAGCTTCAGCAGGAGAAGCGGCAGCTCCGGC 2400

Qy   2341 aggagctcgagagcctcatgaaggagcaggacctgctggagaccaagctcaggtcctacg 2400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2401 AGGAGCTCGAGAGCCTCATGAAGGAGCAGGACCTGCTGGAGACCAAGCTCAGGTCCTACG 2460

Qy   2401 agagggagaagaccagcttcggccccgcgctggaggagacccagtgggaggtgaggccac 2460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2461 AGAGGGAGAAGACCAGCTTCGGCCCCGCGCTGGAGGAGACCCAGTGGGAGGTGAGGCCAC 2520

Qy   2461 acagggctcatgggtttgggtggtcagcggtttggcgccagtacccccctctccttctgg 2520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2521 ACAGGGCTCATGGGTTTGGGTGGTCAGCGGTTTGGCGCCAGTACCCCCCTCTCCTTCTGG 2580

Qy   2521 tgctggccaatagcgtgcaaacacagaccgcgcaggcaagcggggctaatgtgctggctt 2580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2581 TGCTGGCCAATAGCGTGCAAACACAGACCGCGCAGGCAAGCGGGGCTAATGTGCTGGCTT 2640

Qy   2581 tatcacccaaagaaggggctccctgcaaaccatgttgggggatcgacttacatctgagct 2640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2641 TATCACCCAAAGAAGGGGCTCCCTGCAAACCATGTTGGGGGATCGACTTACATCTGAGCT 2700

Qy   2641 tcctcctgtccccaccatcaccctcatggctcctagatttcagtttcccaagtgagccat 2700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2701 TCCTCCTGTCCCCACCATCACCCTCATGGCTCCTAGATTTCAGTTTCCCAAGTGAGCCAT 2760

Qy   2701 taaatcatgaagccggaagccagatgaccaaggcccagccaggctgtgggctgacctccc 2760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2761 TAAATCATGAAGCCGGAAGCCAGATGACCAAGGCCCAGCCAGGCTGTGGGCTGACCTCCC 2820

Qy   2761 ttccatcagctcccaggaggctcagaagaagaacaagccgtgcctgagttcaggcggggc 2820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2821 TTCCATCAGCTCCCAGGAGGCTCAGAAGAAGAACAAGCCGTGCCTGAGTTCAGGCGGGGC 2880

Qy   2821 caggggcccaagagagcacagaatgcatttgttgctttggagggagggactgcacccact 2880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2881 CAGGGGCCCAAGAGAGCACAGAATGCATTTGTTGCTTTGGAGGGAGGGACTGCACCCACT 2940

Qy   2881 agtaagagggaccctattggtggcaggtttcagtgatggaagtggccactccttgctgaa 2940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   2941 AGTAAGAGGGACCCTATTGGTGGCAGGTTTCAGTGATGGAAGTGGCCACTCCTTGCTGAA 3000

Qy   2941 gtgtaagtggaacttctatttggtgagctgagatggaaacctaggagaggaagtaaagag 3000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3001 GTGTAAGTGGAACTTCTATTTGGTGAGCTGAGATGGAAACCTAGGAGAGGAAGTAAAGAG 3060

Qy   3001 tcccccactcacacacttacacactcacacacactcactcacccggtcacacgtggaaat 3060
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3061 TCCCCCACTCACACACTTACACACTCACACACACTCACTCACCCGGTCACACGTGGAAAT 3120
```

```
Qy   3061 gaggcatctgtacctgaccgtgctggagaaccccataacctctgcatctattagtgggaa 3120
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3121 GAGGCATCTGTACCTGACCGTGCTGGAGAACCCCATAACCTCTGCATCTATTAGTGGGAA 3180

Qy   3121 agcagcttttctcaccagcctggtggtctggatgactcatggagttcaagcccatcgttg 3180
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3181 AGCAGCTTTTCTCACCAGCCTGGTGGTCTGGATGACTCATGGAGTTCAAGCCCATCGTTG 3240

Qy   3181 aggctctttacatgctcgcacccagcttggtctgtccacgtgcctgcctcacccccagtt 3240
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3241 AGGCTCTTTACATGCTCGCACCCAGCTTGGTCTGTCCACGTGCCTGCCTCACCCCCAGTT 3300

Qy   3241 cagagtccaaatctcagtctacacgcaaacccctggctatgtgcaagtcaacaaccagtg 3300
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3301 CAGAGTCCAAATCTCAGTCTACACGCAAACCCCTGGCTATGTGCAAGTCAACAACCAGTG 3360

Qy   3301 gtttaacttgcccactgctggcagctgtatcacccccatttaacaccaatggtattggtt 3360
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3361 GTTTAACTTGCCCACTGCTGGCAGCTGTATCACCCCCATTTAACACCAATGGTATTGGTT 3420

Qy   3361 ttggtgtcagcctgatttctgtcatcgatgtttatgcccacatcctctgacctcacccct 3420
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3421 TTGGTGTCAGCCTGATTTCTGTCATCGATGTTTATGCCCACATCCTCTGACCTCACCCCT 3480

Qy   3421 gcatgcacccagccctcctctctcctgtctactggagtaaagactacctcacaaattcac 3480
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3481 GCATGCACCCAGCCCTCCTCTCTCCTGTCTACTGGAGTAAAGACTACCTCACAAATTCAC 3540

Qy   3481 tgctgtacccagtgactagtatcatgctggcttggatgcagagcccaatccacatctgtc 3540
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3541 TGCTGTACCCAGTGACTAGTATCATGCTGGCTTGGATGCAGAGCCCAATCCACATCTGTC 3600

Qy   3541 aaacgaggaatcattttcttctcctcttgctcttctttctctatttcccacccctatccc 3600
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3601 AAACGAGGAATCATTTTCTTCTCCTCTTGCTCTTCTTTCTCTATTTCCCACCCCTATCCC 3660

Qy   3601 ccatcaaaatttggccaagagcaatgatgaaaaccgaagccacaggttagacccatgtgt 3660
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3661 CCATCAAAATTTGGCCAAGAGCAATGATGAAAACCGAAGCCACAGGTTAGACCCATGTGT 3720

Qy   3661 ctctggatcttggccatctggggtcatgggagaccaaggccagtctggctgaatcttaag 3720
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3721 CTCTGGATCTTGGCCATCTGGGGTCATGGGAGACCAAGGCCAGTCTGGCTGAATCTTAAG 3780

Qy   3721 agtgaatgaagtccagagcatgtggctctacagaatggattcttggaactagcctggaag 3780
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3781 AGTGAATGAAGTCCAGAGCATGTGGCTCTACAGAATGGATTCTTGGAACTAGCCTGGAAG 3840

Qy   3781 ccaccttcacatttcctttcacagtagaaatttccccttgccctcagtgaaacactgcac 3840
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3841 CCACCTTCACATTTCCTTTCACAGTAGAAATTTCCCCTTGCCCTCAGTGAAACACTGCAC 3900

Qy   3841 agtcctggagaaaatccgaccctacccaggatgcgtgcttgggaccaagaatttcattcc 3900
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3901 AGTCCTGGAGAAAATCCGACCCTACCCAGGATGCGTGCTTGGGACCAAGAATTTCATTCC 3960

Qy   3901 aaggccaaccctgtattcatgccacgaagggagtgacacagtcatggctgaggcatgggc 3960
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3961 AAGGCCAACCCTGTATTCATGCCACGAAGGGAGTGACACAGTCATGGCTGAGGCATGGGC 4020

Qy   3961 ctggctttgaacctcagcttgaccacttatgatccaggtgattgtaaatacattagccat 4020
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4021 CTGGCTTTGAACCTCAGCTTGACCACTTATGATCCAGGTGATTGTAAATACATTAGCCAT 4080

Qy   4021 ggtggcaatggggtatagtgattaaactgttgggatcaaatctctactcttatacttttat 4080
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4081 GGTGGCAATGGGGTATAGTGATTAAACTGTTGGGATCAAATCTCTACTCTTATACTTTAT 4140

Qy   4081 attttatatatatatatatatatatatatattagccctcaggctggtcacttcac 4140
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4141 ATTTTATATATATATATATATAATATATATATATTAGCCCTCAGGCTGGTCACTTCAC 4200
```

```
Qy   4141 cagctgtttgctatcataacctctctgtgcctcagtttcattgatgtaaattgaggacta 4200
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4201 CAGCTGTTTGCTATCATAACCTCTCTGTGCCTCAGTTTCATTGATGTAAATTGAGGACTA 4260

Qy   4201 ctaatagtacctacttcatcgggttgtaaggaatagatgagcaaatgtatggcttggcac 4260
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4261 CTAATAGTACCTACTTCATCGGGTTGTAAGGAATAGATGAGCAAATGTATGGCTTGGCAC 4320

Qy   4261 ttaataacactacaaattattagtgaaagtatgtttataataatatacttctgtgtggct 4320
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4321 TTAATAACACTACAAATTATTAGTGAAAGTATGTTTATAATAATATACTTCTGTGTGGCT 4380

Qy   4321 aggcgtggtggctcacgcctgcaatcccagcactttgggaggcagaggcaggcagagcac 4380
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4381 AGGCGTGGTGGCTCACGCCTGCAATCCCAGCACTTTGGGAGGCAGAGGCAGGCAGAGCAC 4440

Qy   4381 ttgaggtcaggaattcgagatcagcctggccaacatgaggaaaccccgtctctactaaaa 4440
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4441 TTGAGGTCAGGAATTCGAGATCAGCCTGGCCAACATGAGGAAACCCCGTCTCTACTAAAA 4500

Qy   4441 atacaaaaatcagccaggcatggtggcaggtgtctgtaatcccagctacttgggaggctg 4500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4501 ATACAAAAATCAGCCAGGCATGGTGGCAGGTGTCTGTAATCCCAGCTACTTGGGAGGCTG 4560

Qy   4501 aggcaggagaatcagaggggaggcggaggttgcagtgagccaagatcacgccactacacc 4560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4561 AGGCAGGAGAATCAGAGGGGAGGCGGAGGTTGCAGTGAGCCAAGATCACGCCACTACACC 4620

Qy   4561 ccagcctaggtgacaaagcgagacttctcaaatattaacaataataatatactatgtgtc 4620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4621 CCAGCCTAGGTGACAAAGCGAGACTTCTCAAATATTAACAATAATAATATACTATGTGTC 4680

Qy   4621 attatacatgatgattattattttatcattttactatatagcctagctcgataacctggg 4680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4681 ATTATACATGATGATTATTATTTTATCATTTTACTATATAGCCTAGCTCGATAACCTGGG 4740

Qy   4681 araaaggtcacagcaatgttcagcttactttcagattggacaaaggctggaatgcctaac 4740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4741 ARAAAGGTCACAGCAATGTTCAGCTTACTTTCAGATTGGACAAAGGCTGGAATGCCTAAC 4800

Qy   4741 accgggccaccgcatccggagtggcttggttattttaggcagctgagctgtcacttccct 4800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4801 ACCGGGCCACCGCATCCGGAGTGGCTTGGTTATTTTAGGCAGCTGAGCTGTCACTTCCCT 4860

Qy   4801 gggtaaggacactcacctcttggcactctgtctccaccccaccctcggcaggtgtgccag 4860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4861 GGGTAAGGACACTCACCTCTTGGCACTCTGTCTCCACCCCACCCTCGGCAGGTGTGCCAG 4920

Qy   4861 aagtcaggcgagatctccctcctgaagcagcagctgaaggagtcccagacggaggtgaac 4920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4921 AAGTCAGGCGAGATCTCCCTCCTGAAGCAGCAGCTGAAGGAGTCCCAGACGGAGGTGAAC 4980

Qy   4921 gccaaggctagcgagatcctgggtctcaaggcacagctgaaggacacgcggggcaagctg 4980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4981 GCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGACACGCGGGGCAAGCTG 5040

Qy   4981 gagggcctggagctgaggacccaggacctggagggcgcccctgcgcaccaagggcctggag 5040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5041 GAGGGCCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGCACCAAGGGCCTGGAG 5100

Qy   5041 ctggaggtctgtgagaatgagctgcagcgcaagaagaacgaggcggagctgctgcgggag 5100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5101 CTGGAGGTCTGTGAGAATGAGCTGCAGCGCAAGAAGAACGAGGCGGAGCTGCTGCGGGAG 5160

Qy   5101 aaggtgaacctgctggagcaggagctgcaggagctgcgggcccaggccgccctggccgc 5160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5161 AAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCCCAGGCCGCCCTGGCCCGC 5220

Qy   5161 gacatggggccgcccaccttccccgaggacgtccctgccctgcagcgggagctggagcgg 5220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5221 GACATGGGGCCGCCCACCTTCCCCGAGGACGTCCCTGCCCTGCAGCGGGAGCTGGAGCGG 5280

Qy   5221 ctgcgggccgagctgcgggaggagcggcaaggccatgaccagatgtcctcgggcttccag 5280
```

```
Db   5281 CTGCGGGCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACCAGATGTCCTCGGGCTTCCAG 5340
Qy   5281 catgagcggctcgtgtggaaggaggagaaggagaaggtgattcagtaccagaaacagctg 5340
Db   5341 CATGAGCGGCTCGTGTGGAAGGAGGAGAAGGAGAAGGTGATTCAGTACCAGAAACAGCTG 5400
Qy   5341 cagcagagctacgtggccatgtaccagcggaaccagcgcctggagaaggccctgcagcag 5400
Db   5401 CAGCAGAGCTACGTGGCCATGTACCAGCGGAACCAGCGCCTGGAGAAGGCCCTGCAGCAG 5460
Qy   5401 ctggcacgtggggacagcgccggggagcccttggaggttgacctggaaggggctgacatc 5460
Db   5461 CTGGCACGTGGGGACAGCGCCGGGGAGCCCTTGGAGGTTGACCTGGAAGGGGCTGACATC 5520
Qy   5461 ccctacgaggacatcatagccactgagatctgaggggctgcctgggaaggcgagtctggg 5520
Db   5521 CCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGCCTGGGAAGGCGAGTCTGGG 5580
Qy   5521 gacctggcactgggaggcagggctctcccgtgcatccccccctgctcagcaattcagaccc 5580
Db   5581 GACCTGGCACTGGGAGGCAGGGCTCTCCCGTGCATCCCCCCTGCTCAGCAATTCAGACCC 5640
Qy   5581 ctctgagagacgccactccctgggacacagacccaggacccccgaggggagggcaggatg 5640
Db   5641 CTCTGAGAGACGCCACTCCCTGGGACACAGACCCAGGACCCCCGAGGGGAGGGCAGGATG 5700
Qy   5641 gcctttccttccctctctgatgtcccagtgctcaccagccctgcagcccaccagacgtca 5700
Db   5701 GCCTTTCCTTCCCTCTCTGATGTCCCAGTGCTCACCAGCCCTGCAGCCCACCAGACGTCA 5760
Qy   5701 ggccctgactcctctggctttcccaggagatgggtccaggggtctgtctgctttggttaa 5760
Db   5761 GGCCCTGACTCCTCTGGCTTTCCCAGGAGATGGGTCCAGGGGTCTGTCTGCTTTGGTTAA 5820
Qy   5761 gggctccctaaactttggcctttgttcgaaatagatatcctctccccctcctccagggaa 5820
Db   5821 GGGCTCCCTAAACTTTGGCCTTTGTTCGAAATAGATATCCTCTCCCCCTCCTCCAGGGAA 5880
Qy   5821 ggtggccacagcaagaacagcggctcccctccgcttctcatcccaacctctttttcctcc 5880
Db   5881 GGTGGCCACAGCAAGAACAGCGGCTCCCCTCCGCTTCTCATCCCAACCTCTTTTTCCTCC 5940
Qy   5881 tggacacattggaatgccttggaaatagaaagaagccatatatgaccagaagccttggaa 5940
Db   5941 TGGACACATTGGAATGCCTTGGAAATAGAAAGAAGCCATATATGACCAGAAGCCTTGGAA 6000
Qy   5941 ccagccccatcagaacctgagctatttttcctctggccgcagaggtgtaggggtggaatga 6000
Db   6001 CCAGCCCCATCAGAACCTGAGCTATTTTCCTCTGGCCGCAGAGGTGTAGGGGTGGAATGA 6060
Qy   6001 gccgcggggaagctggctttgaaacctcagggctgtcccagccccggcaagccacaggaa 6060
Db   6061 GCCGCGGGGAAGCTGGCTTTGAAACCTCAGGGCTGTCCCAGCCCCGGCAAGCCACAGGAA 6120
Qy   6061 ggaggggagagacaggcagcccagcagtgtggagaccctgccacagccagaggagggcag 6120
Db   6121 GGAGGGGAGAGACAGGCAGCCCAGCAGTGTGGAGACCCTGCCACAGCCAGAGGAGGGCAG 6180
Qy   6121 agggagaatccaagggttgagagccagtggcgggtgatggccagcccctggggcccagcc 6180
Db   6181 AGGGAGAATCCAAGGGTTGAGAGCCAGTGGCGGGTGATGGCCAGCCCCTGGGGCCCAGCC 6240
Qy   6181 cctgtttactggttcttgcaaatgggagctgagcagcctctggacagccagtgaccttg 6240
Db   6241 CCTGTTTACTGGTTCTTGCAAATGGGAGCTGAGCAGCCTCTGGACAGCCAGTGACCTTTG 6300
Qy   6241 acctcggtgaccactcttctttaagccatagaccctgaggccctgggctgggtgctggga 6300
Db   6301 ACCTCGGTGACCACTCTTCTTTAAGCCATAGACCCTGAGGCCCTGGGCTGGGTGCTGGGA 6360
Qy   6301 agggagggttgaaaccaccgtgaaccagagggtgtggctttccagkcaccctcagggagc 6360
Db   6361 AGGGAGGGTTGAAACCACCGTGAACCAGAGGGTGTGGCTTTCCAGKCACCCTCAGGGAGC 6420
Qy   6361 ctccccatctgtccagctggggccagaggctgggagtccctacctgcttcacgttggccg 6420
Db   6421 CTCCCCATCTGTCCAGCTGGGGCCAGAGGCTGGGAGTCCCTACCTGCTTCACGTTGGCCG 6480
Qy   6421 gcggctactctggaatgtttttccctccccagaatcaagctttgcttgatccagaagag 6480
Db   6481 GCGGCTACTCTGGAATGTTTTCCCTCCCCAGAATCAAGCTTTTGCTTGATCCAGAAGAG 6540
Qy   6481 cccatatcactaagatggcatatatgtgatctgggcattttcctcctctgcctacagcca 6540
Db   6541 CCCATATCACTAAGATGGCATATATGTGATCTGGGCATTTTCCTCCTCTGCCTACAGCCA 6600
Qy   6541 ggtttagcggcaaaccttcccccttagcaccttcagggctgagttctgggtttctagag 6600
Db   6601 GGTTTAGCGGCAAACCTTTCCCCCTTAGCACCTTCAGGGCTGAGTTCTGGGTTTCTAGAG 6660
Qy   6601 gtcaggacggctcctcagagcgccaggaagccagagccccaagcaggacgaaaaagaggc 6660
Db   6661 GTCAGGACGGCTCCTCAGAGCGCCAGGAAGCCAGAGCCCCAAGCAGGACGAAAAAGAGGC 6720
Qy   6661 atacacacagcagtgtgaatagcctggccaccagccatcctccctccacctcaagacccc 6720
Db   6721 ATACACACAGCAGTGTGAATAGCCTGGCCACCAGCCATCCTCCCTCCACCTCAAGACCCC 6780
Qy   6721 catttgtccsagactaaaggatccagagagcagctccctttctcaggagcttgggcagtg 6780
Db   6781 CATTTGTCCSAGACTAAAGGATCCAGAGAGCAGCTCCCTTTCTCAGGAGCTTGGGCAGTG 6840
Qy   6781 ccccagggagtccagggtttctctgcagatgtgcggagcgggaggcggtggtagagagag 6840
Db   6841 CCCCAGGGAGTCCAGGGTTTCTCTGCAGATGTGCGGAGCGGGAGGCGGTGGTAGAGAGAG 6900
Qy   6841 ataaaaggtggagtttctctgttgtttggttcagggattttatttttaatttttatgagac 6900
Db   6901 ATAAAAGGTGGAGTTTCTCTGTTGTTTGGTTCAGGGATTTTATTTTTAATTTTATGAGAC 6960
Qy   6901 agggtcttgctctgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcc 6960
Db   6961 AGGGTCTTGCTCTGTCCCCCAGGCTGGAGTGCAGTGGCATGATCATAGCTCACTGCAGCC 7020
Qy   6961 tcatactcctgggctcaagcaatcctcctgcctcagccttccaactagctgggactacag 7020
Db   7021 TCATACTCCTGGGCTCAAGCAATCCTCCTGCCTCAGCCTTCCAACTAGCTGGGACTACAG 7080
Qy   7021 gtgcgcgccaccgtgcctggctaacttttcatttttttgtagggacggggtctcgtttt 7080
Db   7081 GTGCGCGCCACCGTGCCTGGCTAACTTTTCATTTTTTTTGTAGGGACGGGGTCTCGTTTT 7140
Qy   7081 gttgccaaagctggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaa 7140
Db   7141 GTTGCCAAAGCTGGTCTCAAACTTGTGGCCTCAAGCAATCCACCTGCCTTGGCCTCCCAA 7200
Qy   7141 agtgctgagattgcagatgtgagccaccgtgcctggccagatttttcttttattcttctt 7200
Db   7201 AGTGCTGAGATTGCAGATGTGAGCCACCGTGCCTGGCCAGATTTTTCTTTTATTCTTCTT 7260
Qy   7201 tctttttctttttgctttcttgtcttttcagaagcaagccagacctagcaggctgttcc 7260
Db   7261 TCTTTTTCTTTTTTGCTTTCTTGTCTTTTCAGAAGCAAGCCAGACCTAGCAGGCTGTTCC 7320
Qy   7261 atgttctatttttgactgtagccacagctgctgttctcaggacagcatcccttcccacat 7320
Db   7321 ATGTTCTATTTTTGACTGTAGCCACAGCTGCTGTTCTCAGGACAGCATCCCTTCCCACAT 7380
Qy   7321 gcctgcgcctgctgcctgctgagatgaggaggggagcgtctgggaacttgcgagtccaag 7380
Db   7381 GCCTGCGCCTGCTGCCTGCTGAGATGAGGAGGGGAGCGTCTGGGAACTTGCGAGTCCAAG 7440
Qy   7381 gccagtccccatttctgcctcgctcaccgctggcccttagagaccccgaggtaggggtgg 7440
Db   7441 GCCAGTCCCCATTTCTGCCTCGCTCACCGCTGGCCCTTAGAGACCCCGAGGTAGGGGTGG 7500
```

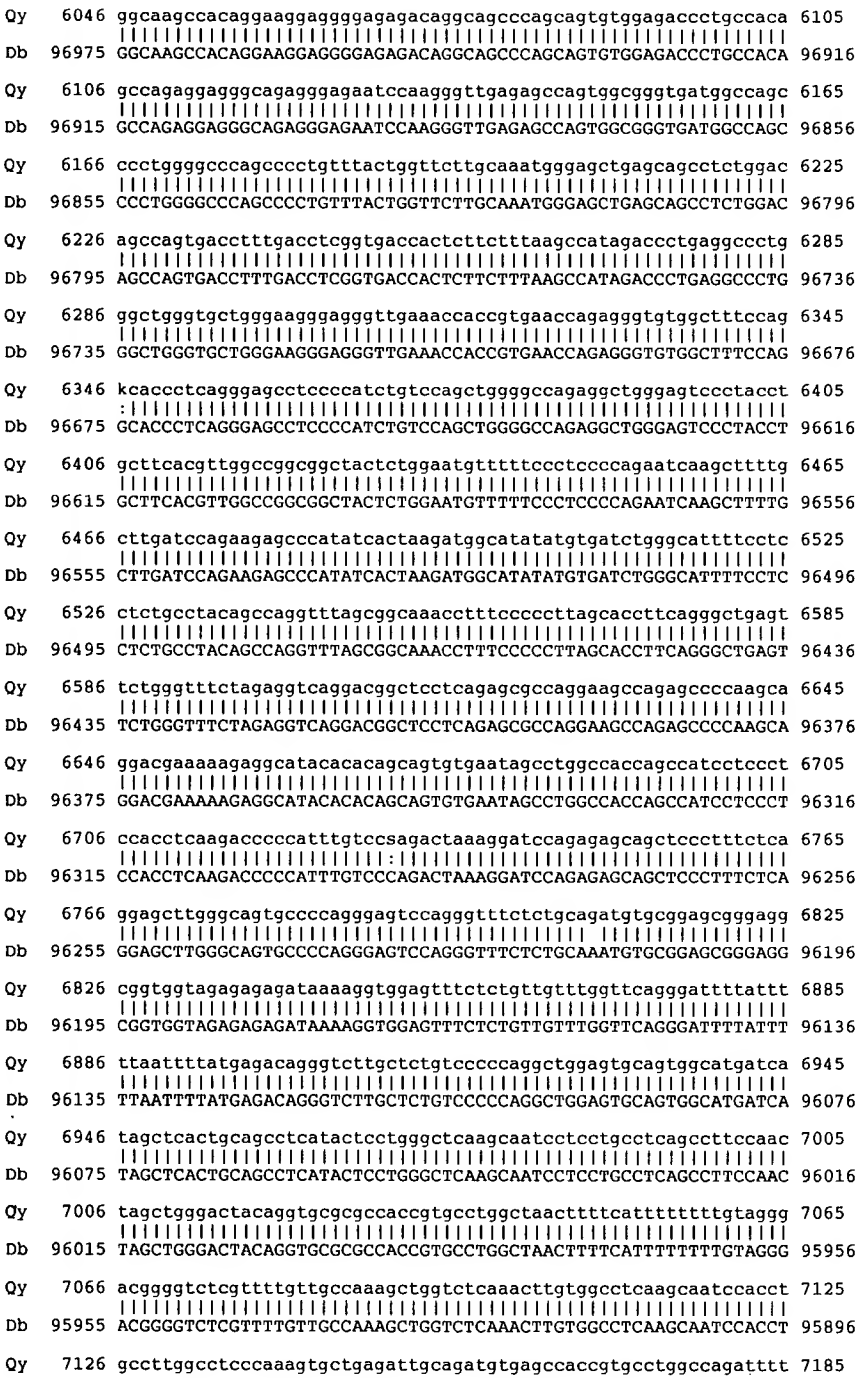
```
Qy   7441 ggagatgcttctctccttgccccccgccctcatgggtcctagcccttccctgagtgcggg 7500
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7501 GGAGATGCTTCTCTCCTTGCCCCCCGCCCTCATGGGTCCTAGCCCTTCCCTGAGTGCGGG 7560

Qy   7501 ctgaggccagagtcacctttttctgtggctggctctaccttcctgtccctgaggttaaacg 7560
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7561 CTGAGGCCAGAGTCACCTTTTCTGTGGCTGGCTCTACCTTCCTGTCCCTGAGGTTAAACG 7620

Qy   7561 gtgcccatcctgccatcctcaaacgacagaggagcttttctggaatttcaaaccattgct 7620
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7621 GTGCCCATCCTGCCATCCTCAAACGACAGAGGAGCTTTTCTGGAATTTCAAACCATTGCT 7680

Qy   7621 cttagtcccaagctaggcttaaacctggaatctacaagccaaaagtccctccctgcctga 7680
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7681 CTTAGTCCCAAGCTAGGCTTAAACCTGGAATCTACAAGCCAAAAGTCCCTCCCTGCCTGA 7740

Qy   7681 gggcagtaccctccattgggcacagtccagacccaagtcaaagatgccccattccttgcg 7740
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7741 GGGCAGTACCCTCCATTGGGCACAGTCCAGACCCAAGTCAAAGATGCCCCATTCCTTGCG 7800

Qy   7741 cctcagccctcagttccttcatttccaccaggccgtgccttgtttgagtttttcctccca 7800
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7801 CCTCAGCCCTCAGTTCCTTCATTTCCACCAGGCCGTGCCTTGTTTGAGTTTTTCCTCCCA 7860

Qy   7801 gtgagactgccccacggagacagaggaaagggctggctccccctccccaggctggagacc 7860
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7861 GTGAGACTGCCCCACGGAGACAGAGGAAAGGGCTGGCTCCCCCTCCCCAGGCTGGAGACC 7920

Qy   7861 ccccccaactccaggaaagagcagtcagagtccagtgctctgcctcagacgttgcctgag 7920
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7921 CCCCCCAACTCCAGGAAAGAGCAGTCAGAGTCCAGTGCTCTGCCTCAGACGTTGCCTGAG 7980

Qy   7921 aagaagtggctgccacacccaggggaaggccctgaggcggaggctgtgctccgccatggt 7980
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   7981 AAGAAGTGGCTGCCACACCCAGGGGAAGGCCCTGAGGCGGAGGCTGTGCTCCGCCATGGT 8040

Qy   7981 gtcccggtaccttccatacacagaggagtgcagccttctccatatctccatggccctgtc 8040
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8041 GTCCCGGTACCTTCCATACACAGAGGAGTGCAGCCTTCTCCATATCTCCATGGCCCTGTC 8100

Qy   8041 ccaggccggcccagatgtgtccccccccaggccttgtcctacgtccaaggtggcagatgtc 8100
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8101 CCAGGCCGGCCCAGATGTGTCCCCCCCAGGCCTTGTCCTACGTCCAAGGTGGCAGATGTC 8160

Qy   8101 ttccctgggctgccaccagccccgccccagagtggcccaccgtggcactagaatgcaag 8160
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8161 TTCCCTGGGCTGCCACCAGCCCCCGCCCCAGAGTGGCCCACCGTGGCACTAGAATGCAAG 8220

Qy   8161 tatcctgcgaccttgcaacctcaccttcctgtgggtgttctttcctgccctgtccaaaag 8220
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8221 TATCCTGCGACCTTGCAACCTCACCTTCCTGTGGGTGTTCTTTCCTGCCCTGTCCAAAAG 8280

Qy   8221 cgccctcactattcttggaccatgccagattctgcctctctggaaagaggctctggacag 8280
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8281 CGCCCTCACTATTCTTGGACCATGCCAGATTCTGCCTCTCTGGAAAGAGGCTCTGGACAG 8340

Qy   8281 cagaagcctccaagcacagagcctggccccaggccccagacagggtgggcttcctgccct 8340
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8341 CAGAAGCCTCCAAGCACAGAGCCTGGCCCCAGGCCCCAGACAGGGTGGGCTTCCTGCCCT 8400

Qy   8341 tccctctgggcacgcctgctggccgacccactgacccactcggatggaccaacctgctct 8400
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8401 TCCCTCTGGGCACGCCTGCTGGCCGACCCACTGACCCACTCGGATGGACCAACCTGCTCT 8460

Qy   8401 gtccccaaaggacgcctgcaggagagagcagcactccgcatcacctcaccaaggatcgga 8460
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8461 GTCCCCAAAGGACGCCTGCAGGAGAGAGCAGCACTCCGCATCACCTCACCAAGGATCGGA 8520

Qy   8461 ctctgcccctggacctgggaacgactggactgtcacggggttccctcctagctctcccag 8520
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8521 CTCTGCCCCTGGACCTGGGAACGACTGGACTGTCACGGGGTTCCCTCCTAGCTCTCCCAG 8580
```

```
Qy   8521 tgaactcctgccaggcacacacagcccctatagcactgagctcacatgggactgggatat 8580
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8581 TGAACTCCTGCCAGGCACACACAGCCCCTATAGCACTGAGCTCACATGGGACTGGGATAT 8640

Qy   8581 gggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggccccagtctgg 8640
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8641 GGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCCCAGTCTGG 8700

Qy   8641 gccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggaggagcagca 8700
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8701 GCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAGGAGCAGCA 8760

Qy   8701 gcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgctgccatga 8760
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8761 GCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGCTGCCATGA 8820

Qy   8761 cagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaaggtggcctt 8820
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8821 CAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGCAAGGTGGCCTT 8880

Qy   8821 tgtgctgcgcctgcaggcagagctggagcccccagcagaggcaggctgggacggaccagc 8880
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8881 TGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGACGGACCAGC 8940

Qy   8881 atctggaagatgtacatagttatttttctctttgtggtttcttgtttggtttggtttgct 8940
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   8941 ATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCT 9000

Qy   8941 tttgacagcttcattttatttttgacgtcacttttttggccatgtaaactatttgtggcaa 9000
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   9001 TTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATTTGTGGCAA 9060

Qy   9001 ttttatgtttttatttatgaataaagaatgccatttctcacgccctct 9048
          |||||||||||||||||||||||||||||||||||||||||||||||
Db   9061 TTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACGCCCTCT 9108
```

```
RESULT   2
AC025853/c
LOCUS       AC025853   172162 bp    DNA             HTG       26-MAY-2000
DEFINITION  Homo sapiens chromosome 8 clone RP11-353K12 map 8, WORKING DRAFT
            SEQUENCE, 25 unordered pieces.
ACCESSION   AC025853
VERSION     AC025853.3  GI:8077000
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 172162)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 8, clone RP11-353K12
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 172162)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
```

```
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7342114.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L7454
                Center clone name: 353_K_12
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 156061 bases at least Q40
                Consensus quality: 163284 bases at least Q30
                Consensus quality: 166861 bases at least Q20
                Insert size: 182000; agarose-fp
                Insert size: 169762; sum-of-contigs
                Quality coverage: 4.0 in Q20 bases; agarose-fp
                Quality coverage: 4.3 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 25 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1104: contig of 1104 bp in length
            *     1105 1204: gap of      100 bp
            *     1205     3226: contig of 2022 bp in length
            *     3227 3326: gap of      100 bp
            *     3327     4615: contig of 1289 bp in length
            *     4616 4715: gap of      100 bp
            *     4716     6570: contig of 1855 bp in length
            *     6571 6670: gap of      100 bp
            *     6671     8592: contig of 1922 bp in length
            *     8593 8692: gap of      100 bp
            *     8693    11929: contig of 3237 bp in length
            *    11930 12029: gap of      100 bp
            *    12030    15549: contig of 3520 bp in length
            *    15550 15649: gap of      100 bp
            *    15650    19338: contig of 3689 bp in length
            *    19339 19438: gap of      100 bp
            *    19439    23177: contig of 3739 bp in length
            *    23178 23277: gap of      100 bp
            *    23278    29032: contig of 5755 bp in length
            *    29033 29132: gap of      100 bp
            *    29133    34839: contig of 5707 bp in length
            *    34840 34939: gap of      100 bp
            *    34940    40597: contig of 5658 bp in length
            *    40598 40697: gap of      100 bp
            *    40698    47278: contig of 6581 bp in length
            *    47279 47378: gap of      100 bp
            *    47379    54796: contig of 7418 bp in length
            *    54797 54896: gap of      100 bp
            *    54897    59726: contig of 4830 bp in length
            *    59727 59826: gap of      100 bp
            *    59827    68926: contig of 9100 bp in length
            *    68927 69026: gap of      100 bp
            *    69027    75947: contig of 6921 bp in length
            *    75948 76047: gap of      100 bp
            *    76048    83423: contig of 7376 bp in length
            *    83424 83523: gap of      100 bp
            *    83524    90514: contig of 6991 bp in length
            *    90515 90614: gap of      100 bp
            *    90615    99034: contig of 8420 bp in length
            *    99035 99134: gap of      100 bp
            *    99135   109649: contig of 10515 bp in length
            *   109650 109749: gap of      100 bp
            *   109750   122299: contig of 12550 bp in length
            *   122300 122399: gap of      100 bp
            *   122400   137369: contig of 14970 bp in length
            *   137370 137469: gap of      100 bp
            *   137470   154199: contig of 16730 bp in length
            *   154200 154299: gap of      100 bp
            *   154300   172162: contig of 17863 bp in length.
FEATURES             Location/Qualifiers
     source          1. .172162
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8"
                     /clone="RP11-353K12"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1104
                     /note="assembly_fragment"
     misc_feature    1205. .3226
                     /note="assembly_fragment"
     misc_feature    3327. .4615
                     /note="assembly_fragment"
     misc_feature    4716. .6570
                     /note="assembly_fragment"
     misc_feature    6671. .8592
                     /note="assembly_fragment"
     misc_feature    8693. .11929
                     /note="assembly_fragment"
     misc_feature    12030. .15549
                     /note="assembly_fragment"
     misc_feature    15650. .19338
                     /note="assembly_fragment"
     misc_feature    19439. .23177
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    23278. .29032
                     /note="assembly_fragment"
     misc_feature    29133. .34839
                     /note="assembly_fragment"
     misc_feature    34940. .40597
                     /note="assembly_fragment"
     misc_feature    40698. .47278
                     /note="assembly_fragment"
     misc_feature    47379. .54796
                     /note="assembly_fragment"
     misc_feature    54897. .59726
                     /note="assembly_fragment"
     misc_feature    59827. .68926
                     /note="assembly_fragment"
     misc_feature    69027. .75947
                     /note="assembly_fragment"
     misc_feature    76048. .83423
                     /note="assembly_fragment"
     misc_feature    83524. .90514
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    90615. .99034
                     /note="assembly_fragment"
     misc_feature    99135. .109649
                     /note="assembly_fragment"
     misc_feature    109750. .122299
                     /note="assembly_fragment"
     misc_feature    122400. .137369
                     /note="assembly_fragment"
     misc_feature    137470. .154199
                     /note="assembly_fragment"
```

```
     misc_feature    154300. .172162
                     /note="assembly_fragment"
BASE COUNT    47201 a  37955 c  38816 g  45788 t   2402 others
ORIGIN

  Query Match             54.5%;  Score 4926.8;  DB 64;  Length 172162;
  Best Local Similarity   99.1%;  Pred. No. 0;
  Matches 4999;  Conservative    3;  Mismatches   20;  Indels   22;  Gaps    4;

Qy   4025 gcaatggggtatagtgattaaactgttgggatcaaatctctactcttatactttatatttt 4084
          ||||||||| |||| |||||||||| ||||||||||||||||||||||   ||| | |||
Db  99013 GCAATGGGGAATAG-GATTAAACTG-TGGGATCAAATCTCTACTCTTAACTTTTTTTTTT 98956

Qy   4085 tatatatatatatatata------------------atatatatatatattagccctca 4125
          ||||||||||||||||||                  ||||||||||||||||||||||
Db  98955 TATATATATATATATATATATATATATATATATATATATATATATATATATTAGCCCTCA 98896

Qy   4126 ggctggtcacttcaccagctgtttgctatcataacctctctgtgcctcagtttcattgat 4185
          ||||||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db  98895 GGCTGGTCACTTCACCAGCTATTTGCTATCATAACCTCTCTGTGCCTCAGTTTCATTGAT 98836

Qy   4186 gtaaattgaggactactaatagtacctacttcatcgggttgtaaggaatagatgagcaaa 4245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98835 GTAAATTGAGGACTACTAATAGTACCTACTTCATCGGGTTGTAAGGAATAGATGAGCAAA 98776

Qy   4246 tgtatggcttggcacttaataacactacaaattattagtgaaagtatgtttataataata 4305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98775 TGTATGGCTTGGCACTTAATAACACTACAAATTATTAGTGAAAGTATGTTTATAATAATA 98716

Qy   4306 tacttctgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgggaggcag 4365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98715 TACTTCTGTGTGGCTAGGCGTGGTGGCTCACGCCTGCAATCCCAGCACTTTGGGAGGCAG 98656

Qy   4366 aggcaggcagagcacttgaggtcaggaattcgagatcagcctggccaacatgaggaaacc 4425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98655 AGGCAGGCAGAGCACTTGAGGTCAGGAATTCGAGATCAGCCTGGCCAACATGAGGAAACC 98596

Qy   4426 ccgtctctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgtaatcccag 4485
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98595 CCGTCTCTACTAAAAATACAAAAATCAGCCAGGCATGGTGGCAGGTGTCTGTAATCCCAG 98536

Qy   4486 ctacttgggaggctgaggcaggagaatcagaggggaggcggaggttgcagtgagccaaga 4545
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98535 CTACTTGGGAGGCTGAGGCAGGAGAATCAGAGGGGAGGCGGAGGTTGCAGTGAGCCAAGA 98476

Qy   4546 tcacgccactacaccccagcctaggtgacaaagcgagacttctcaaatattaacaataat 4605
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Db  98475 TCACGCCACTACACCCCAGCCTAGGTGACAAAGCGAGACTTCTCAAATAATAACAATAAT 98416

Qy   4606 aatatactatgtgtcattatacatgatgattattatttttatcattttactatatagccta 4665
          |||||||||||||||||||||||| |||||||||||||||||||||||||||||||||||
Db  98415 AATATACTATGTGTCATTATACATGATTATTATTATTTTATCATTTTACTATATAGCCTA 98356

Qy   4666 gctcgataacctgggaraaaggtcacagcaatgttcagcttactttcagattggacaaag 4725
          ||||||||||||||| :|||||||||||||||||||||||||||||||||||||||||||
Db  98355 GCTCGATAACCTGGGGAAAAGGTCACAGCAATGTTCAGCTTACTTTCAGATTGGACAAAG 98296

Qy   4726 gctggaatgcctaacaccgggccaccgcatccggagtggcttggttattttaggcagctg 4785
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98295 GCTGGAATGCCTAACACCGGGCCACCGCATCCGGAGTGGCTTGGTTATTTTAGGCAGCTG 98236

Qy   4786 agctgtcacttccctgggtaaggacactcacctcttggcactctgtctccaccccaccct 4845
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98235 AGCTGTCACTTCCCTGGGTAAGGACACTCACCTCTTGGCACTCTGTCTCCACCCCACCCT 98176

Qy   4846 cggcaggtgtgccagaagtcaggcgagatctccctcctgaagcagcagctgaaggagtcc 4905
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98175 CGGCAGGTGTGCCAGAAGTCAGGCGAGATCTCCCTCCTGAAGCAGCAGCTGAAGGAGTCC 98116

Qy   4906 cagacggaggtgaacgccaaggctagcgagatcctgggtctcaaggcacagctgaaggac 4965
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db. 98115 CAGACGGAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAAGGAC 98056

Qy   4966 acgcggggcaagctggagggcctggagctgaggacccaggacctggagggcgccctgcgc 5025
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  98055 ACGCGGGGCAAGCTGGAGGGCCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCTGCGC 97996

Qy   5026 accaagggcctggagctggaggtctgtgagaatgagctgcagcgcaagaagaacgaggcg 5085
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97995 ACCAAGGGCCTGGAGCTGGAGGTCTGTGAGAATGAGCTGCAGCGCAAGAAGAACGAGGCG 97936

Qy   5086 gagctgctgcgggagaaggtgaacctgctggagcaggagctgcaggagctgcgggcccag 5145
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97935 GAGCTGCTGCGGGAGAAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGCCCAG 97876

Qy   5146 gccgccctggcccgcgacatggggccgcccaccttccccgaggacgtccctgccctgcag 5205
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97875 GCCGCCCTGGCCCGCGACATGGGGCCGCCCACCTTCCCCGAGGACGTCCCTGCCCTGCAG 97816

Qy   5206 cgggagctggagcggctgcgggccgagctgcgggaggagcggcaaggccatgaccagatg 5265
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97815 CGGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACCAGATG 97756

Qy   5266 tcctcgggcttccagcatgagcggctcgtgtggaaggaggagaaggagaaggtgattcag 5325
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97755 TCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAAGGAGAAGGTGATTCAG 97696

Qy   5326 taccagaaacagctgcagcagagctacgtggccatgtaccagcggaaccagcgcctggag 5385
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97695 TACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGCGGAACCAGCGCCTGGAG 97636

Qy   5386 aaggccctgcagcagctggcacgtggggacagcgccggggagcccttggaggttgacctg 5445
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97635 AAGGCCCTGCAGCAGCTGGCACGTGGGGACAGCGCCGGGGAGCCCTTGGAGGTTGACCTG 97576

Qy   5446 gaaggggctgacatcccctacgaggacatcatagccactgagatctgaggggctgcctgg 5505
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||
Db  97575 GAAGGGGCTGACATCCCCTACGAGGACATCATAGCCACTGAGATTTGAGGGGCTGCCTGG 97516

Qy   5506 gaaggcgagtctggggacctggcactgggaggcagggctctcccgtgcatcccccctgct 5565
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97515 GAAGGCGAGTCTGGGGACCTGGCACTGGGAGGCAGGGCTCTCCCGTGCATCCCCCCTGCT 97456

Qy   5566 cagcaattcagacccctctgagagacgccactccctgggacacagacccaggacccccga 5625
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db  97455 CAGCAATTCAGACCCCTTTGAGAGACGCCACTCCCTGGGACACAGACCCAGGACCCCCGA 97396

Qy   5626 ggggagggcaggatggccttccttccctctctgatgtcccagtgctcaccagccctgca 5685
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97395 GGGGAGGGCAGGATGGCCTTTCCTTCCCTCTCTGATGTCCCAGTGCTCACCAGCCCTGCA 97336

Qy   5686 gcccaccagacgtcaggccctgactcctctggctttcccaggagatgggtccaggggtct 5745
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97335 GCCCACCAGACGTCAGGCCCTGACTCCTCTGGCTTTCCCAGGAGATGGGTCCAGGGGTCT 97276

Qy   5746 gtctgctttggttaagggctccctaaactttggcctttgttcgaaatagatatcctctcc 5805
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97275 GTCTGCTTTGGTTAAGGGCTCCCTAAACTTTGGCCTTTGTTCGAAATAGATATCCTCTCC 97216

Qy   5806 ccctcctccagggaaggtggccacagcaagaacagcggctcccctccgcttctcatccca 5865
          |||||||||||||||||||||||||||||||| ||||||||||||||| ||||||||||||
Db  97215 CCCTCCTCCAGGGAAGGTGGCCACAGCAAGTACAGCGGCTCCCCTCTGCTTCTCATCCCA 97156

Qy   5866 acctctttttcctcctggacacattggaatgccttggaaatagaaagaagccatatatga 5925
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97155 ACCTCTTTTTCCTCCTGGACACATTGGAATGCCTTGGAAATAGAAAGAAGCCATATATGA 97096

Qy   5926 ccagaagccttggaaccagccccatcagaacctgagctatttttcctctggccgcagaggt 5985
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97095 CCAGAAGCCTTGGAACCAGCCCCATCAGAACCTGAGCTATTTTCCTCTGGCCGCAGAGGT 97036

Qy   5986 gtaggggtggaatgagccgcggggaagctggctttgaaacctcagggctgtcccagcccc 6045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  97035 GTAGGGGTGGAATGAGCCGCGGGGAAGCTGGCTTTGAAACCTCAGGGCTGTCCCAGCCCC 96976
```


Qy 6046 ggcaagccacaggaaggaggggagagacaggcagcccagcagtgtggagaccctgccaca 6105
Db 96975 GGCAAGCCACAGGAAGGAGGGGAGAGACAGGCAGCCCAGCAGTGTGGAGACCCTGCCACA 96916
Qy 6106 gccagaggagggcagagggagaatccaagggttgagagccagtggcgggtgatggccagc 6165
Db 96915 GCCAGAGGAGGGCAGAGGGAGAATCCAAGGGTTGAGAGCCAGTGGCGGGTGATGGCCAGC 96856
Qy 6166 ccctggggcccagcccctgtttactggttcttgcaaatgggagctgagcagcctctggac 6225
Db 96855 CCCTGGGGCCCAGCCCCTGTTTACTGGTTCTTGCAAATGGGAGCTGAGCAGCCTCTGGAC 96796
Qy 6226 agccagtgaccttgacctcggtgaccactcttctttaagccatagaccctgaggccctg 6285
Db 96795 AGCCAGTGACCTTTGACCTCGGTGACCACTCTTCTTTAAGCCATAGACCCTGAGGCCCTG 96736
Qy 6286 ggctgggtgctgggaagggagggttgaaaccaccgtgaaccagagggtgtggctttccag 6345
Db 96735 GGCTGGGTGCTGGGAAGGGAGGGTTGAAACCACCGTGAACCAGAGGGTGTGGCTTTCCAG 96676
Qy 6346 kcaccctcagggagcctccccatctgtccagctggggccagaggctgggagtccctacct 6405
Db 96675 GCACCCTCAGGGAGCCTCCCCATCTGTCCAGCTGGGGCCAGAGGCTGGGAGTCCCTACCT 96616
Qy 6406 gcttcacgttggccggcggctactctggaatgtttttccctccccagaatcaagcttttg 6465
Db 96615 GCTTCACGTTGGCCGGCGGCTACTCTGGAATGTTTTTCCCTCCCCAGAATCAAGCTTTTG 96556
Qy 6466 cttgatccagaagagcccatatcactaagatggcatatatgtgatctgggcattttcctc 6525
Db 96555 CTTGATCCAGAAGAGCCCATATCACTAAGATGGCATATATGTGATCTGGGCATTTTCCTC 96496
Qy 6526 ctctgcctacagccaggtttagcggcaaaccttcccccttagcaccttcagggctgagt 6585
Db 96495 CTCTGCCTACAGCCAGGTTTAGCGGCAAACCTTTCCCCCTTAGCACCTTCAGGGCTGAGT 96436
Qy 6586 tctgggtttctagaggtcaggacggctcctcagagcgccaggaagccagagccccaagca 6645
Db 96435 TCTGGGTTTCTAGAGGTCAGGACGGCTCCTCAGAGCGCCAGGAAGCCAGAGCCCCAAGCA 96376
Qy 6646 ggacgaaaaagaggcatacacacagcagtgtgaatagcctggccaccagccatcctccct 6705
Db 96375 GGACGAAAAAGAGGCATACACACAGCAGTGTGAATAGCCTGGCCACCAGCCATCCTCCCT 96316
Qy 6706 ccacctcaagacccccatttgtccsagactaaaggatccagagagcagctccctttctca 6765
Db 96315 CCACCTCAAGACCCCCATTTGTCCCAGACTAAAGGATCCAGAGAGCAGCTCCCTTTCTCA 96256
Qy 6766 ggagcttgggcagtgccccagggagtccagggtttctctgcagatgtgcggagcgggagg 6825
Db 96255 GGAGCTTGGGCAGTGCCCCAGGGAGTCCAGGGTTTCTCTGCAAATGTGCGGAGCGGGAGG 96196
Qy 6826 cggtggtagagagagatataaaaggtggagtttctctgttgtttggttcagggattttattt 6885
Db 96195 CGGTGGTAGAGAGAGATAAAAGGTGGAGTTTCTCTGTTGTTTGGTTCAGGGATTTTATTT 96136
Qy 6886 ttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtggcatgatca 6945
Db 96135 TTAATTTTATGAGACAGGGTCTTGCTCTGTCCCCCAGGCTGGAGTGCAGTGGCATGATCA 96076
Qy 6946 tagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagccttccaac 7005
Db 96075 TAGCTCACTGCAGCCTCATACTCCTGGGCTCAAGCAATCCTCCTGCCTCAGCCTTCCAAC 96016
Qy 7006 tagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttttttgtaggg 7065
Db 96015 TAGCTGGGACTACAGGTGCGCGCCACCGTGCCTGGCTAACTTTTCATTTTTTTTTGTAGGG 95956
Qy 7066 acggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagcaatccacct 7125
Db 95955 ACGGGGTCTCGTTTTGTTGCCAAAGCTGGTCTCAAACTTGTGGCCTCAAGCAATCCACCT 95896
Qy 7126 gccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggccagatttt 7185
Db 95895 GCCTTGGCCTCCCAAAGTGCTGAGATTGCAGATGTGAGCCACCGTGCCTGGCCAGATTTT 95836
Qy 7186 tcttttattcttctttcttttttcttttttgctttcttgtcttttcagaagcaagccagac 7245
Db 95835 TCTTTTATTCTTCTTTCTTTTTCTTTTTTGCTTTCTTGTCTTTTCAGAAGCAAGCCAGAC 95776
Qy 7246 ctagcaggctgttccatgttctatttttgactgtagccacagctgctgttctcaggacag 7305
Db 95775 CTAGCAGGCTGTTCCATGTTCTATTTTTGACTGTAGCCACAGCTGCTGTTCTCAGGACAG 95716
Qy 7306 catcccttcccacatgcctgcgcctgctgcctgctgagatgaggaggggagcgtctggga 7365
Db 95715 CATCCCTTCCCACATGCCTGCGCCTGCTGCCTGCTGAGATGAGGAGGGGAGCGTCTGGGA 95656
Qy 7366 acttgcgagtccaaggccagtccccatttctgcctcgctcaccgctggcccttagagacc 7425
Db 95655 ACTTGCGAGTCCAAGGCCAGTCCCCATTTCTGCCTCGCTCACCGCTGGCCCTTAGAGACC 95596
Qy 7426 ccgaggtaggggtggggagatgcttctctccttgccccccgccctcatgggtcctagccc 7485
Db 95595 CCGAGGTAGGGGTGGGGAGATGCTTCTCTCCTTGCCCCCCGCCCTCATGGGTCCTAGCCC 95536
Qy 7486 ttccctgagtgcgggctgaggccagagtcacctttttctgtggctggctctaccttcctgt 7545
Db 95535 TTCCCTGAGTGCGGGCTGAGGCCAGAGTCACCTTTTCTGTGGCTGGCTCTACCTTCCTGT 95476
Qy 7546 ccctgaggttaaacggtgcccatcctgccatcctcaaacgacagaggagcttttctggaa 7605
Db 95475 CCCTGAGGTTAAACGGTGCCCATCCTGCCATCCTCAAACGACAGAGGAGCTTTTCTGGAA 95416
Qy 7606 tttcaaaccattgctcttagtcccaagctaggcttaaacctggaatctacaagccaaaag 7665
Db 95415 TTTCAAACCATTGCTCTTAGTCCCAAGCTAGGCTTAAACCTGGAATCTACAAGCCAAAAG 95356
Qy 7666 tccctccctgcctgagggcagtaccctccattgggcacagtccagacccaagtcaaagat 7725
Db 95355 TCCCTCCCTGCCTGAGGGCAGTACCCTCCATTGGGCACAGTCCAGACCCAAGTCAAAGAT 95296
Qy 7726 gccccattccttgcgcctcagccctcagttccttcatttccaccaggccgtgccttgttt 7785
Db 95295 GCCCCATTCCTTGCGCCTCAGCCCTCAGTTCCTTCATTTCCACCAGGCCGTGCCTTGTTT 95236
Qy 7786 gagtttttcctcccagtgagactgccccacggagacagaggaaagggctggctccccctc 7845
Db 95235 GAGTTTTTCCTCCCAGTGAGACTGCCCCACGGAGACAGAGGAAAGGGCTGGCTCCCCCTC 95176
Qy 7846 cccaggctggaga-ccccccccaactccaggaaagagcagtcagagtccagtgctctgcc 7904
Db 95175 CCCAGGCTGGAGACCCCCCCCCAACTCCAGGAAAGAGCAGTCAGAGTCCAGTGCTCTGCC 95116
Qy 7905 tcagacgttgcctgagaagaagtggctgccacacccaggggaaggccctgaggcggaggc 7964
Db 95115 TCAGACGTTGCCTGAGAAGAAGTGGCTGCCACACCCAGGGGAAGGCCCTGAGGGGGGGGG 95056
Qy 7965 tgtgctccgccatggtgtcccggtaccttccatacacagaggagtgcagccttctccata 8024
Db 95055 TGTGCTCCGCCATGGTGTCCCGGTACCTTCCATACACAGAGGAGTGCAGCCTTCTCCATA 94996
Qy 8025 tctccatggccctgtcccaggccggcccagatgtgtcccccccaggccttgtcctacgtc 8084
Db 94995 TCTCCATGGCCCTGTCCCAGGCCGGCCCAGATGTGTCCCCCCCAGGCCTTGTCCTACGTC 94936
Qy 8085 caaggtggcagatgtcttccctgggctgccaccagccccgccccagagtggcccaccgt 8144
Db 94935 CAAGGTGGCAGATGTCTTCCCTGGGCTGCCACCAGCCCCCGCCCCAGAGTGGCCCACCGT 94876
Qy 8145 ggcactagaatgcaagtatcctgcgaccttgcaacctcaccttcctgtgggtgttctttc 8204
Db 94875 GGCACTAGAATGCAAGTATCCTGCGACCTTGCAACCTCACCTTCCTGTGGGTGTTCTTTC 94816
Qy 8205 ctgccctgtccaaaagcgccctcactattcttggaccatgccagattctgcctctctgga 8264

```
Db  94815 CTGCCCTGTCCAAAAGCGCCCTCACTATTCTTGGACCATGCCAGATTCTGCCTCTCTGGA 94756

Qy   8265 aagaggctctggacagcagaagcctccaagcacagagcctggccccaggccccagacagg 8324
          |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94755 AAGAGGGTCTGGACAGCAGAAGCCTCCAAGCACAGAGCCTGGCCCCAGGCCCCAGACAGG 94696

Qy   8325 gtgggcttcctgcccttccctctgggcacgcctgctggccgacccactgacccactcgga 8384
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94695 GTGGGCTTCCTGCCCTTCCCTCTGGGCACGCCTGCTGGCCGACCCACTGACCCACTCGGA 94636

Qy   8385 tggaccaacctgctctgtccccaaaggacgcctgcaggagagagcagcactccgcatcac 8444
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94635 TGGACCAACCTGCTCTGTCCCCAAAGGACGCCTGCAGGAGAGAGCAGCACTCCGCATCAC 94576

Qy   8445 ctcaccaaggatcggactctgcccctggacctgggaacgactggactgtcacggggttcc 8504
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94575 CTCACCAAGGATCGGACTCTGCCCCTGGACCTGGGAACGACTGGACTGTCACGGGGTTCC 94516

Qy   8505 ctcctagctctcccagtgaactcctgccaggcacacacagcccctatagcactgagctca 8564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94515 CTCCTAGCTCTCCCAGTGAACTCCTGCCAGGCACACACAGCCCCTATAGCACTGAGCTCA 94456

Qy   8565 catgggactgggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtg 8624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94455 CATGGGACTGGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTG 94396

Qy   8625 cctggccccagtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcct 8684
          |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db  94395 CCTGGCCCCAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCT 94336

Qy   8685 ggctggaggagcagcagcttggggagagggatttccctgcagacctcaagccatcatgcg 8744
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94335 GGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCG 94276

Qy   8745 gtgggtgctgccatgacagaggctgcacccctgggccagcggggctgctcacccacctct 8804
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94275 GTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCT 94216

Qy   8805 tgtgcaaggtggccctttgtgctgcgcctgcaggcagagctggagccccagcagaggcag 8864
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94215 TGTGCAAGGTGGCCTTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAG 94156

Qy   8865 gctgggacggaccagcatctggaagatgtacatagttatttttctctttgtggtttcttg 8924
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94155 GCTGGGACGGACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTG 94096

Qy   8925 tttggtttggtttgcttttgacagcttcattttatttttgacgtcacttttttggccatgt 8984
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94095 TTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATGT 94036

Qy   8985 aaactatttgtggcaattttatgttttttatttatgaataaagaatgccatttctcacgcc 9044
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  94035 AAACTATTTGTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACGCC 93976

Qy   9045 ctct 9048
          ||||
Db  93975 CTCT 93972


RESULT   3
AF123659
LOCUS       AF123659     5492 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens FEZ1 (FEZ1) mRNA, complete cds.
ACCESSION   AF123659
VERSION     AF123659.1  GI:4572475
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 5492)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 5492)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .5492
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
     gene            1. .5492
                     /gene="FEZ1"
     CDS             112. .1902
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23840.1"
                     /db_xref="GI:4572476"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQRKKNEAELLREKVNLLEQELQELRAQAALARDMGPPTFPEDVPALQRELERLR
                     AELREERQGHDQMSSGFQHERLVWKEEKEKVIQYQKQLQQSYVAMYQRNQRLEKALQQ
                     LARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT     1137 a   1704 c   1565 g   1086 t
ORIGIN


  Query Match           46.4%;  Score 4199.8;  DB 10;  Length 5492;
  Best Local Similarity   99.9%;  Pred. No. 0;
  Matches 4201;  Conservative    2;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   4842 ccctcggcaggtgtgccagaagtcaggcgagatctccctcctgaagcagcagctgaagga 4901
          ||   || ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1251 CCAGTGGGAGGTGTGCCAGAAGTCAGGCGAGATCTCCCTCCTGAAGCAGCAGCTGAAGGA 1310

Qy   4902 gtcccagacggaggtgaacgccaaggctagcgagatcctgggtctcaaggcacagctgaa 4961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1311 GTCCCAGACGGAGGTGAACGCCAAGGCTAGCGAGATCCTGGGTCTCAAGGCACAGCTGAA 1370

Qy   4962 ggacacgcggggcaagctggagggcctggagctgaggacccaggacctggagggcgccct 5021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1371 GGACACGCGGGGCAAGCTGGAGGGCCTGGAGCTGAGGACCCAGGACCTGGAGGGCGCCCT 1430

Qy   5022 gcgcaccaagggcctggagctggaggtctgtgagaatgagctgcagcgcaagaagaacga 5081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1431 GCGCACCAAGGGCCTGGAGCTGGAGGTCTGTGAGAATGAGCTGCAGCGCAAGAAGAACGA 1490

Qy   5082 ggcggagctgctgcgggagaaggtgaacctgctggagcaggagctgcaggagctgcgggc 5141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1491 GGCGGAGCTGCTGCGGGAGAAGGTGAACCTGCTGGAGCAGGAGCTGCAGGAGCTGCGGGC 1550

Qy   5142 ccaggccgccctggcccgcgacatggggccgcccaccttccccgaggacgtccctgccct 5201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1551 CCAGGCCGCCCTGGCCCGCGACATGGGGCCGCCCACCTTCCCCGAGGACGTCCCTGCCCT 1610

Qy   5202 gcagcgggagctggagcggctgcgggccgagctgcgggaggagcggcaaggccatgacca 5261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1611 GCAGCGGGAGCTGGAGCGGCTGCGGGCCGAGCTGCGGGAGGAGCGGCAAGGCCATGACCA 1670

Qy   5262 gatgtcctcgggcttccagcatgagcggctcgtgtggaaggaggagaaggagaaggtgat 5321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db  1671 GATGTCCTCGGGCTTCCAGCATGAGCGGCTCGTGTGGAAGGAGGAGAAGGAGAAGGTGAT 1730

Qy  5322 tcagtaccagaaacagctgcagcagagctacgtggccatgtaccagcggaaccagcgcct 5381
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1731 TCAGTACCAGAAACAGCTGCAGCAGAGCTACGTGGCCATGTACCAGCGGAACCAGCGCCT 1790

Qy  5382 ggagaaggccctgcagcagctggcacgtggggacagcgccggggagcccttggaggttga 5441
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1791 GGAGAAGGCCCTGCAGCAGCTGGCACGTGGGGACAGCGCCGGGGAGCCCTTGGAGGTTGA 1850

Qy  5442 cctggaaggggctgacatcccctacgaggacatcatagccactgagatctgaggggctgc 5501
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1851 CCTGGAAGGGGCTGACATCCCCTACGAGGACATCATAGCCACTGAGATCTGAGGGGCTGC 1910

Qy  5502 ctgggaaggcgagtctggggacctggcactgggaggcagggctctcccgtgcatcccccc 5561
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1911 CTGGGAAGGCGAGTCTGGGGACCTGGCACTGGGAGGCAGGGCTCTCCCGTGCATCCCCCC 1970

Qy  5562 tgctcagcaattcagacccctctgagagacgccactccctgggacacagacccaggaccc 5621
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  1971 TGCTCAGCAATTCAGACCCCTCTGAGAGACGCCACTCCCTGGGACACAGACCCAGGACCC 2030

Qy  5622 ccgaggggagggcaggatggccttccttccctctctgatgtcccagtgctcaccagccc 5681
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2031 CCGAGGGGAGGGCAGGATGGCCTTTCCTTCCCTCTCTGATGTCCCAGTGCTCACCAGCCC 2090

Qy  5682 tgcagcccaccagacgtcaggccctgactcctctggctttcccaggagatgggtccaggg 5741
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2091 TGCAGCCCACCAGACGTCAGGCCCTGACTCCTCTGGCTTTCCCAGGAGATGGGTCCAGGG 2150

Qy  5742 gtctgtctgctttggttaagggctccctaaactttggcctttgttcgaaatagatatcct 5801
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2151 GTCTGTCTGCTTTGGTTAAGGGCTCCCTAAACTTTGGCCTTTGTTCGAAATAGATATCCT 2210

Qy  5802 ctccccctcctccagggaaggtggccacagcaagaacagcggctcccctccgcttctcat 5861
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2211 CTCCCCCTCCTCCAGGGAAGGTGGCCACAGCAAGAACAGCGGCTCCCCTCCGCTTCTCAT 2270

Qy  5862 cccaacctctttttcctcctggacacattggaatgccttggaaatagaaagaagccatat 5921
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2271 CCCAACCTCTTTTTCCTCCTGGACACATTGGAATGCCTTGGAAATAGAAAGAAGCCATAT 2330

Qy  5922 atgaccagaagccttggaaccagccccatcagaacctgagctattttcctctggccgcag 5981
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2331 ATGACCAGAAGCCTTGGAACCAGCCCCATCAGAACCTGAGCTATTTTCCTCTGGCCGCAG 2390

Qy  5982 aggtgtaggggtggaatgagccgcggggaagctggctttgaaacctcagggctgtcccag 6041
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2391 AGGTGTAGGGGTGGAATGAGCCGCGGGGAAGCTGGCTTTGAAACCTCAGGGCTGTCCCAG 2450

Qy  6042 ccccggcaagccacaggaaggaggggagagacaggcagcccagcagtgtggagaccctgc 6101
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2451 CCCCGGCAAGCCACAGGAAGGAGGGGAGAGACAGGCAGCCCAGCAGTGTGGAGACCCTGC 2510

Qy  6102 cacagccagaggagggcagagggagaatccaagggttgagagccagtggcgggtgatggc 6161
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2511 CACAGCCAGAGGAGGGCAGAGGGAGAATCCAAGGGTTGAGAGCCAGTGGCGGGTGATGGC 2570

Qy  6162 cagcccctggggcccagcccctgtttactggttcttgcaaatgggagctgagcagcctct 6221
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2571 CAGCCCCTGGGGCCCAGCCCCTGTTTACTGGTTCTTGCAAATGGGAGCTGAGCAGCCTCT 2630

Qy  6222 ggacagccagtgaccttgacctcggtgaccactcttctttaagccatagaccctgaggc 6281
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2631 GGACAGCCAGTGACCTTTGACCTCGGTGACCACTCTTCTTTAAGCCATAGACCCTGAGGC 2690

Qy  6282 cctgggctgggtgctgggaagggagggttgaaaccaccgtgaaccagagggtgtggcttt 6341
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2691 CCTGGGCTGGGTGCTGGGAAGGGAGGGTTGAAACCACCGTGAACCAGAGGGTGTGGCTTT 2750

Qy  6342 ccagkcaccctcagggagcctccccatctgtccagctggggccagaggctgggagtccct 6401
         ||||:||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2751 CCAGGCACCCTCAGGGAGCCTCCCCATCTGTCCAGCTGGGGCCAGAGGCTGGGAGTCCCT 2810

Qy  6402 acctgcttcacgttggccggcggctactctggaatgtttttccctccccagaatcaagct 6461
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2811 ACCTGCTTCACGTTGGCCGGCGGCTACTCTGGAATGTTTTTCCCTCCCCAGAATCAAGCT 2870

Qy  6462 tttgcttgatccagaagagcccatatcactaagatggcatatatgtgatctgggcatttt 6521
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2871 TTTGCTTGATCCAGAAGAGCCCATATCACTAAGATGGCATATATGTGATCTGGGCATTTT 2930

Qy  6522 cctcctctgcctacagccaggtttagcggcaaacctttccccttagcaccttcagggct 6581
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2931 CCTCCTCTGCCTACAGCCAGGTTTAGCGGCAAACCTTTCCCCCTTAGCACCTTCAGGGCT 2990

Qy  6582 gagttctgggtttctagaggtcaggacggctcctcagagcgccaggaagccagagccccca 6641
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  2991 GAGTTCTGGGTTTCTAGAGGTCAGGACGGCTCCTCAGAGCGCCAGGAAGCCAGAGCCCCA 3050

Qy  6642 agcaggacgaaaaagaggcatacacacagcagtgtgaatagcctggccaccagccatcct 6701
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3051 AGCAGGACGAAAAAGAGGCATACACACAGCAGTGTGAATAGCCTGGCCACCAGCCATCCT 3110

Qy  6702 ccctccacctcaagacccccatttgtccsagactaaaggatccagagagcagctcccttt 6761
         ||||||||||||||||||||||||||||:||||||||||||||||||||||||||||||
Db  3111 CCCTCCACCTCAAGACCCCCATTTGTCCCAGACTAAAGGATCCAGAGAGCAGCTCCCTTT 3170

Qy  6762 ctcaggagcttgggcagtgccccagggagtccagggtttctctgcagatgtgcggagcgg 6821
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3171 CTCAGGAGCTTGGGCAGTGCCCCAGGGAGTCCAGGGTTTCTCTGCAGATGTGCGGAGCGG 3230

Qy  6822 gaggcggtggtagagagagatAAAAGGTGGAGTTTCTCTGTTGTTTGGTTCAGGGATTTT 6881
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3231 GAGGCGGTGGTAGAGAGAGATAAAAGGTGGAGTTTCTCTGTTGTTTGGTTCAGGGATTTT 3290

Qy  6882 atttttaattttatgagacagggtcttgctctgtccccaggctggagtgcagtggcatg 6941
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3291 ATTTTTAATTTTATGAGACAGGGTCTTGCTCTGTCCCCCAGGCTGGAGTGCAGTGGCATG 3350

Qy  6942 atcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagccttc 7001
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3351 ATCATAGCTCACTGCAGCCTCATACTCCTGGGCTCAAGCAATCCTCCTGCCTCAGCCTTC 3410

Qy  7002 caactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttttgt 7061
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3411 CAACTAGCTGGGACTACAGGTGCGCGCCACCGTGCCTGGCTAACTTTTCATTTTTTTTGT 3470

Qy  7062 agggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagcaatcc 7121
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3471 AGGGACGGGGTCTCGTTTTGTTGCCAAAGCTGGTCTCAAACTTGTGGCCTCAAGCAATCC 3530

Qy  7122 acctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggccaga 7181
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3531 ACCTGCCTTGGCCTCCCAAAGTGCTGAGATTGCAGATGTGAGCCACCGTGCCTGGCCAGA 3590

Qy  7182 tttttcttttattcttctttcttttttcttttttgctttcttgtctttttcagaagcaagcc 7241
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3591 TTTTTCTTTTATTCTTCTTTCTTTTTCTTTTTTGCTTTCTTGTCTTTTCAGAAGCAAGCC 3650

Qy  7242 agacctagcaggctgttccatgttctatttttgactgtagccacagctgctgttctcagg 7301
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3651 AGACCTAGCAGGCTGTTCCATGTTCTATTTTTGACTGTAGCCACAGCTGCTGTTCTCAGG 3710

Qy  7302 acagcatcccttcccacatgcctgcgcctgctgcctgctgagatgaggaggggagcgtct 7361
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3711 ACAGCATCCCTTCCCACATGCCTGCGCCTGCTGCCTGCTGAGATGAGGAGGGGAGCGTCT 3770

Qy  7362 gggaacttgcgagtccaaggccagtccccatttctgcctcgctcaccgctggccccttaga 7421
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3771 GGGAACTTGCGAGTCCAAGGCCAGTCCCCATTTCTGCCTCGCTCACCGCTGGCCCTTAGA 3830

Qy  7422 gaccccgaggtaggggtggggagatgcttctctccttgccccccgccctcatgggtccta 7481
         ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db  3831 GACCCCGAGGTAGGGGTGGGGAGATGCTTCTCTCCTTGCCCCCCGCCCTCATGGGTCCTA 3890
```

```
Qy   7482 gcccttccctgagtgcgggctgaggccagagtcacctttttctgtggctggctctaccttc 7541
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3891 GCCCTTCCCTGAGTGCGGGCTGAGGCCAGAGTCACCTTTTCTGTGGCTGGCTCTACCTTC 3950

Qy   7542 ctgtccctgaggttaaacggtgcccatcctgccatcctcaaacgacagaggagctttttct 7601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   3951 CTGTCCCTGAGGTTAAACGGTGCCCATCCTGCCATCCTCAAACGACAGAGGAGCTTTTCT 4010

Qy   7602 ggaatttcaaaccattgctcttagtcccaagctaggcttaaacctggaatctacaagcca 7661
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4011 GGAATTTCAAACCATTGCTCTTAGTCCCAAGCTAGGCTTAAACCTGGAATCTACAAGCCA 4070

Qy   7662 aaagtccctccctgcctgagggcagtaccctccattgggcacagtccagacccaagtcaa 7721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4071 AAAGTCCCTCCCTGCCTGAGGGCAGTACCCTCCATTGGGCACAGTCCAGACCCAAGTCAA 4130

Qy   7722 agatgccccattccttgcgcctcagccctcagttccttcatttccaccaggccgtgccttt 7781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4131 AGATGCCCCATTCCTTGCGCCTCAGCCCTCAGTTCCTTCATTTCCACCAGGCCGTGCCTT 4190

Qy   7782 gtttgagtttttcctcccagtgagactgccccacggagacagaggaaagggctggctccc 7841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4191 GTTTGAGTTTTTCCTCCCAGTGAGACTGCCCCACGGAGACAGAGGAAAGGGCTGGCTCCC 4250

Qy   7842 cctccccaggctggagaccccccccaactccaggaaagagcagtcagagtccagtgctct 7901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4251 CCTCCCCAGGCTGGAGACCCCCCCCAACTCCAGGAAAGAGCAGTCAGAGTCCAGTGCTCT 4310

Qy   7902 gcctcagacgttgcctgagaagaagtggctgccacacccaggggaaggccctgaggcgga 7961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4311 GCCTCAGACGTTGCCTGAGAAGAAGTGGCTGCCACACCCAGGGGAAGGCCCTGAGGCGGA 4370

Qy   7962 ggctgtgctccgccatggtgtcccggtaccttccatacacagaggagtgcagccttctcc 8021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4371 GGCTGTGCTCCGCCATGGTGTCCCGGTACCTTCCATACACAGAGGAGTGCAGCCTTCTCC 4430

Qy   8022 atatctccatggccctgtcccaggccggcccagatgtgtccccccaggccttgtcctac 8081
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4431 ATATCTCCATGGCCCTGTCCCAGGCCGGCCCAGATGTGTCCCCCCCAGGCCTTGTCCTAC 4490

Qy   8082 gtccaaggtggcagatgtcttccctgggctgccaccagcccccgccccagagtggcccac 8141
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4491 GTCCAAGGTGGCAGATGTCTTCCCTGGGCTGCCACCAGCCCCCGCCCCAGAGTGGCCCAC 4550

Qy   8142 cgtggcactagaatgcaagtatcctgcgaccttgcaacctcaccttcctgtgggtgttct 8201
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4551 CGTGGCACTAGAATGCAAGTATCCTGCGACCTTGCAACCTCACCTTCCTGTGGGTGTTCT 4610

Qy   8202 ttcctgccctgtccaaaagcgccctcactattcttggaccatgccagattctgcctctct 8261
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4611 TTCCTGCCCTGTCCAAAAGCGCCCTCACTATTCTTGGACCATGCCAGATTCTGCCTCTCT 4670

Qy   8262 ggaaagaggctctggacagcagaagcctccaagcacagagcctggccccaggccccagac 8321
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4671 GGAAAGAGGCTCTGGACAGCAGAAGCCTCCAAGCACAGAGCCTGGCCCCAGGCCCCAGAC 4730

Qy   8322 agggtgggcttcctgcccttccctctgggcacgcctgctggccgacccactgacccactc 8381
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4731 AGGGTGGGCTTCCTGCCCTTCCCTCTGGGCACGCCTGCTGGCCGACCCACTGACCCACTC 4790

Qy   8382 ggatggaccaacctgctctgtccccaaaggacgcctgcaggagagagcagcactccgcat 8441
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4791 GGATGGACCAACCTGCTCTGTCCCCAAAGGACGCCTGCAGGAGAGAGCAGCACTCCGCAT 4850

Qy   8442 cacctcaccaaggatcggactctgcccctggacctgggaacgactggactgtcacggggt 8501
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4851 CACCTCACCAAGGATCGGACTCTGCCCCTGGACCTGGGAACGACTGGACTGTCACGGGGT 4910

Qy   8502 tccctcctagctctcccagtgaactcctgccaggcacacacagcccctatagcactgagc 8561
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4911 TCCCTCCTAGCTCTCCCAGTGAACTCCTGCCAGGCACACACAGCCCCTATAGCACTGAGC 4970

Qy.  8562 tcacatgggactgggatatgggggcatctcttccccagagaggcactcagtgagcctcct 8621
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   4971 TCACATGGGACTGGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCT 5030

Qy   8622 gtgcctggccccagtctgggccatctcttaggtgagacagttgcccgaaactaagccagg 8681
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5031 GTGCCTGGCCCCAGTCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGG 5090

Qy   8682 cctggctggaggagcagcagcttggggagagggatttccctgcagacctcaagccatcat 8741
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5091 CCTGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCAT 5150

Qy   8742 gcggtgggtgctgccatgacagaggctgcacccctgggccagcggggctgctcacccacc 8801
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5151 GCGGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACC 5210

Qy   8802 tcttgtgcaaggtggcctttgtgctgcgcctgcaggcagagctggagcccccagcagagg 8861
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5211 TCTTGTGCAAGGTGGCCTTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGG 5270

Qy   8862 caggctgggacggaccagcatctggaagatgtacatagttatttttctctttgtggtttc 8921
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5271 CAGGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTC 5330

Qy   8922 ttgtttggtttggtttgcttttgacagcttcattttatttttgacgtcacttttttggcca 8981
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5331 TTGTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCA 5390

Qy   8982 tgtaaactatttgtggcaatttttatgtttttatttatgaataaagaatgccatttctcac 9041
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   5391 TGTAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCAC 5450

Qy   9042 gccctct 9048
          |||||||
Db   5451 GCCCTCT 5457
```

```
RESULT   4
AF123656
LOCUS       AF123656     1515 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens clone D14 FEZ1 (FEZ1) mRNA, alternatively spliced,
            complete cds.
ACCESSION   AF123656
VERSION     AF123656.1  GI:4572469
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1515)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 1515)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1515
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
                     /clone="D14"
     gene            1. .1515
                     /gene="FEZ1"
     CDS             1. .1515
                     /gene="FEZ1"
                     /codon_start=1
```

```
                     /product="FEZ1"
                     /protein_id="AAD23837.1"
                     /db_xref="GI:4572470"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQQSYVAMYQRNQRLEKALQQLARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT      336 a    483 c    482 g    214 t
ORIGIN


  Query Match             8.9%;  Score 807.2;  DB 10;  Length 1515;
  Best Local Similarity   99.6%;  Pred. No. 2.6e-119;
  Matches  809;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 GGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGG 404

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGA 464

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 524

Qy   1826 ccggaactccatgtccagcctgcccacacacagcaccagcagcagctaccagctggaccc 1885
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCC 584

Qy   1886 gctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcaccca 1945
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585 GCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCA 644

Qy   1946 gggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccgacgg 2005
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 GGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGG 704

Qy   2006 aggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctcccc 2065
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCC 764

Qy   2066 catctccacggacgagtgcagcatccaggagctggaacagaagctgttggagagggaggg 2125
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGG 824

Qy   2126 cgccctccagaagctgcagcgcagctttgaggagaaggagcttgcctccagcctggccta 2185
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGGCCTA 884

Qy   2186 cgaggagcggccgcggcgctgcagggacgagctggagggcccggagcccaaaggcggcaa 2245
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 CGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAA 944

Qy   2246 caagctcaagcaggcctcgcagaagagccagcgcgcgcagcaggtcctgcacctgcaggt 2305
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    945 CAAGCTCAAGCAGGCCTCGCAGAAGAGCCAGCGCGCGCAGCAGGTCCTGCACCTGCAGGT 1004

Qy   2306 actgcagcttcagcaggagaagcggcagctccggcaggagctcgagagcctcatgaagga 2365
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 ACTGCAGCTTCAGCAGGAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGA 1064

Qy   2366 gcaggacctgctggagaccaagctcaggtcctacgagagggagaagaccagcttcggccc 2425
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 GCAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGGAGAAGACCAGCTTCGGCCC 1124

Qy   2426 cgcgctggaggagacccagtgggaggtgaggc 2457
           |||||||||||||||||||||||||||||| | |
Db   1125 CGCGCTGGAGGAGACCCAGTGGGAGGTGTGCC 1156


RESULT   5
AF123655
LOCUS       AF123655     1614 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens clone T8D145M4 FEZ1 (FEZ1) mRNA, alternatively
            spliced, complete cds.
ACCESSION   AF123655
VERSION     AF123655.1  GI:4572467
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1614)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 1614)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1614
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
                     /clone="T8D145M4"
     gene            1. .1614
                     /gene="FEZ1"
     CDS             1. .1614
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23836.1"
                     /db_xref="GI:4572468"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQRKKNEAELLREKHERLVWKEEKEKVIQYQKQLQQSYVAMYQRNQRLEKALQQL
                     ARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT      367 a    500 c    522 g    225 t
ORIGIN


  Query Match             8.9%;  Score 807.2;  DB 10;  Length 1614;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-119;
  Matches  809;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 GGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGG 404

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGA 464

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 524
```

```
Qy   1826 ccggaactccatgtccagcctgcccacacacagcaccagcagcagctaccagctggaccc 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCC 584

Qy   1886 gctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcaccca 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585 GCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCA 644

Qy   1946 gggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccgacgg 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 GGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGG 704

Qy   2006 aggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctcccc 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCC 764

Qy   2066 catctccacggacgagtgcagcatccaggagctggaacagaagctgttggagagggaggg 2125
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGG 824

Qy   2126 cgccctccagaagctgcagcgcagctttgaggagaaggagcttgcctccagcctggccta 2185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGGCCTA 884

Qy   2186 cgaggagcggccgcggcgctgcagggacgagctggagggcccggagcccaaaggcggcaa 2245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 CGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAA 944

Qy   2246 caagctcaagcaggcctcgcagaagagccagcgcgcgcagcaggtcctgcacctgcaggt 2305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    945 CAAGCTCAAGCAGGCCTCGCAGAAGAGCCAGCGCGCGCAGCAGGTCCTGCACCTGCAGGT 1004

Qy   2306 actgcagcttcagcaggagaagcggcagctccggcaggagctcgagagcctcatgaagga 2365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 ACTGCAGCTTCAGCAGGAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGA 1064

Qy   2366 gcaggacctgctggagaccaagctcaggtcctacgagagggagaagaccagcttcggccc 2425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 GCAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGGAGAAGACCAGCTTCGGCCC 1124

Qy   2426 cgcgctggaggagacccagtgggaggtgaggc 2457
          ||||||||||||||||||||||||||| | |
Db   1125 CGCGCTGGAGGAGACCCAGTGGGAGGTGTGCC 1156
```

```
RESULT   6
AF123657
LOCUS       AF123657     1692 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens clone G3611 FEZ1 (FEZ1) mRNA, alternatively spliced,
            complete cds.
ACCESSION   AF123657
VERSION     AF123657.1  GI:4572471
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1692)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 1692)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1692
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
                     /clone="G3611"
     gene            1. .1692
                     /gene="FEZ1"
     CDS             1. .1692
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23838.1"
                     /db_xref="GI:4572472"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQRKKNEAELLREKVNLLERLRAELREERQGHDQMSSGFQHERLVWKEEKEKVIQ
                     YQKQLQQSYVAMYQRNQRLEKALQQLARGDSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT      380 a    522 c    554 g    236 t
ORIGIN
```

```
  Query Match             8.9%;  Score 807.2;  DB 10;  Length 1692;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-119;
  Matches  809;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 GGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGG 404

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGA 464

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 524

Qy   1826 ccggaactccatgtccagcctgcccacacacagcaccagcagcagctaccagctggaccc 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCC 584

Qy   1886 gctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcaccca 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585 GCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCA 644

Qy   1946 gggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccgacgg 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 GGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGG 704

Qy   2006 aggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctcccc 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCC 764

Qy   2066 catctccacggacgagtgcagcatccaggagctggaacagaagctgttggagagggaggg 2125
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db    765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGG 824

Qy   2126 cgccctccagaagctgcagcgcagctttgaggagaaggagcttgcctccagcctggccta 2185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGGCCTA 884

Qy   2186 cgaggagcggccgcggcgctgcagggacgagctggagggcccggagcccaaaggcggcaa 2245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 CGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAA 944

Qy   2246 caagctcaagcaggcctcgcagaagagccagcgcgcgcagcaggtcctgcacctgcaggt 2305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db    945 CAAGCTCAAGCAGGCCTCGCAGAAGAGCCAGCGCGCGCAGCAGGTCCTGCACCTGCAGGT 1004

Qy   2306 actgcagcttcagcaggagaagcggcagctccggcaggagctcgagagcctcatgaagga 2365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 ACTGCAGCTTCAGCAGGAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGA 1064

Qy   2366 gcaggacctgctggagaccaagctcaggtcctacgagagggagaagaccagcttcggccc 2425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 GCAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGGAGAAGACCAGCTTCGGCCC 1124

Qy   2426 cgcgctggaggagacccagtgggaggtgaggc 2457
          |||||||||||||||||||||||||||||| | |
Db   1125 CGCGCTGGAGGAGACCCAGTGGGAGGTGTGCC 1156


RESULT   7
AF123658
LOCUS       AF123658     1722 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens clone G3612 FEZ1 (FEZ1) mRNA, alternatively spliced,
            complete cds.
ACCESSION   AF123658
VERSION     AF123658.1  GI:4572473
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 1722)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 1722)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .1722
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
                     /clone="G3612"
     gene            1. .1722
                     /gene="FEZ1"
     CDS             1. .1722
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23839.1"
                     /db_xref="GI:4572474"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSTSSSYQLDPLVTPVGPTSRFGGSAHNITQGIV
                     LQDSNMMSLKALSFSDGGSKLGHSNKADKGPSCVRSPISTDECSIQELEQKLLEREGA
                     LQKLQRSFEEKELASSLAYEERPRRCRDELEGPEPKGGNKLKQASQKSQRAQQVLHLQ
                     VLQLQQEKRQLRQELESLMKEQDLLETKLRSYEREKTSFGPALEETQWEVCQKSGEIS
                     LLKQQLKESQTEVNAKASEILGLKAQLKDTRGKLEGLELRTQDLEGALRTKGLELEVC
                     ENELQRKKNEAELLREKVNLLEQELQELRAQAALARDMGPPTFPEDVPALQRELERLV
                     WKEEKEKVIQYQKQLQQSYVAMYQRNQRLEKALQQLARGDSAGEPLEVDLEGADIPYE
                     DIIATEI"
BASE COUNT      381 a    541 c    563 g    237 t
ORIGIN


  Query Match             8.9%;  Score 807.2;  DB 10;  Length 1722;
  Best Local Similarity   99.6%;  Pred. No. 2.5e-119;
  Matches  809;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    345 GGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGG 404

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    405 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGA 464

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    465 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 524

Qy   1826 ccggaactccatgtccagcctgcccacacacagcaccagcagcagctaccagctggaccc 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    525 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCC 584

Qy   1886 gctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcaccca 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    585 GCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCA 644

Qy   1946 gggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccgacgg 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    645 GGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGG 704

Qy   2006 aggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctcccc 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    705 AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCC 764

Qy   2066 catctccacggacgagtgcagcatccaggagctggaacagaagctgttggagagggaggg 2125
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Db    765 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGG 824

Qy   2126 cgccctccagaagctgcagcgcagctttgaggagaaggagcttgcctccagcctggccta 2185
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    825 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGGCCTA 884

Qy   2186 cgaggagcggccgcggcgctgcagggacgagctggagggcccggagcccaaaggcggcaa 2245
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    885 CGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGCCCAAAGGCGGCAA 944

Qy   2246 caagctcaagcaggcctcgcagaagagccagcgcgcgcagcaggtcctgcacctgcaggt 2305
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    945 CAAGCTCAAGCAGGCCTCGCAGAAGAGCCAGCGCGCGCAGCAGGTCCTGCACCTGCAGGT 1004

Qy   2306 actgcagcttcagcaggagaagcggcagctccggcaggagctcgagagcctcatgaagga 2365
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1005 ACTGCAGCTTCAGCAGGAGAAGCGGCAGCTCCGGCAGGAGCTCGAGAGCCTCATGAAGGA 1064

Qy   2366 gcaggacctgctggagaccaagctcaggtcctacgagagggagaagaccagcttcggccc 2425
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db   1065 GCAGGACCTGCTGGAGACCAAGCTCAGGTCCTACGAGAGGGAGAAGACCAGCTTCGGCCC 1124

Qy   2426 cgcgctggaggagacccagtgggaggtgaggc 2457
          |||||||||||||||||||||||||||||| | |
Db   1125 CGCGCTGGAGGAGACCCAGTGGGAGGTGTGCC 1156


RESULT   8
AF123654
LOCUS       AF123654      633 bp    mRNA            PRI       07-APR-1999
DEFINITION  Homo sapiens clone E264162 FEZ1 (FEZ1) mRNA, alternatively spliced,
            complete cds.
ACCESSION   AF123654
VERSION     AF123654.1  GI:4572465
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 633)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
```

```
  TITLE     The FEZ1 gene at chromosome 8p22 encodes a leucine-zipper protein,
            and its expression is altered in multiple human tumors
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 96 (7), 3928-3933 (1999)
  MEDLINE   99199287
REFERENCE   2  (bases 1 to 633)
  AUTHORS   Ishii,H., Baffa,R., Numata,S.I., Murakumo,Y., Rattan,S., Inoue,H.,
            Mori,M., Fidanza,V., Alder,H. and Croce,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-JAN-1999) Immunology/Microbiology, Kimmel Cancer
            Institute, 233S 10th street, Philadelphia, PA 19107, USA
FEATURES             Location/Qualifiers
     source          1. .633
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="8"
                     /map="8p22"
                     /clone="E264162"
     gene            1. .633
                     /gene="FEZ1"
     CDS             1. .633
                     /gene="FEZ1"
                     /codon_start=1
                     /product="FEZ1"
                     /protein_id="AAD23835.1"
                     /db_xref="GI:4572466"
                     /translation="MGSVSSLISGHSFHSKHCRASQYKLRKSSHLKKLNRYSDGLLRF
                     GFSQDSGHGKSSSKMGKSEDFFYIKVSQKARGSHHPDYTALSSGDLGGQAGVDFDPST
                     PPKLMPFSNQLEMGSEKGAVRPTAFKPVLPRSGAILHSSPESASHQLHPAPPDKPKEQ
                     ELKPGLCSGALSDSGRNSMSSLPTHSAGEPLEVDLEGADIPYEDIIATEI"
BASE COUNT      136 a    217 c    175 g    105 t
ORIGIN
```

```
  Query Match             3.8%;  Score 346;  DB 10;  Length 633;
  Best Local Similarity   100.0%;  Pred. No. 6.4e-46;
  Matches  346;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      52 atgggcagcgtcagtagcctcatctccggccacagcttccacagcaagcactgccgggct 111
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       1 ATGGGCAGCGTCAGTAGCCTCATCTCCGGCCACAGCTTCCACAGCAAGCACTGCCGGGCT 60

Qy     112 tcgcagtacaagctgcgcaagtcctcccacctcaagaagctcaaccggtattccgacggg 171
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      61 TCGCAGTACAAGCTGCGCAAGTCCTCCCACCTCAAGAAGCTCAACCGGTATTCCGACGGG 120

Qy     172 ctgctgaggtttggcttctcccaggactccggtcacggcaagtccagctccaaaatgggc 231
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     121 CTGCTGAGGTTTGGCTTCTCCCAGGACTCCGGTCACGGCAAGTCCAGCTCCAAAATGGGC 180

Qy     232 aagagcgaagacttcttctacatcaaggtcagccagaaagcccggggctcccatcaccca 291
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     181 AAGAGCGAAGACTTCTTCTACATCAAGGTCAGCCAGAAAGCCCGGGGCTCCCATCACCCA 240

Qy     292 gattacacggcactgtccagcggggatttaggggccaggctggggtggactttgacccg 351
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     241 GATTACACGGCACTGTCCAGCGGGGATTTAGGGGGCCAGGCTGGGGTGGACTTTGACCCG 300

Qy     352 tccacacccccaagctcatgcccttctccaatcagctagaaatgg 397
           |||||||||||||||||||||||||||||||||||||||||||||
Db     301 TCCACACCCCCCAAGCTCATGCCCTTCTCCAATCAGCTAGAAATGG 346
```

```
RESULT   9
AC025412/c
LOCUS       AC025412   163683 bp    DNA             HTG       26-MAY-2000
DEFINITION  Homo sapiens chromosome 1 clone RP11-306I4 map 1, WORKING DRAFT
            SEQUENCE, 22 unordered pieces.
ACCESSION   AC025412
VERSION     AC025412.3  GI:8077107
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 163683)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens chromosome 1, clone RP11-306I4
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 163683)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G.,
            Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
            Collymore,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S.,
            Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,
            Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
            Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
            Klein,J., LaRocque,K., Lamazares,R., Landers,T., Lehoczky,J.,
            Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Marquis,N.,
            McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheeters,R.,
            Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
            Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Olivar,T.M., Oliver,J., Peterson,K., Pierre,N.,
            Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
            Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
            Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
            Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
            Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
            Young,G., Zainoun,J., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (08-MAR-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On May 25, 2000 this sequence version replaced gi:7342167.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L8132
                Center clone name: 306_I_4
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 100% of reads
                Chemistry: Dye-terminator Big Dye; 100% of reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 152913 bases at least Q40
                Consensus quality: 158568 bases at least Q30
                Consensus quality: 160537 bases at least Q20
                Insert size: 170000; agarose-fp
                Insert size: 161583; sum-of-contigs
                Quality coverage: 4.2 in Q20 bases; agarose-fp
                Quality coverage: 4.5 in Q20 bases; sum-of-contigs
            --------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 22 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     1333: contig of 1333 bp in length
            *    1334 1433: gap of      100 bp
            *    1434     2643: contig of 1210 bp in length
            *    2644 2743: gap of      100 bp
            *    2744     4634: contig of 1891 bp in length
            *    4635 4734: gap of      100 bp
            *    4735     7179: contig of 2445 bp in length
            *    7180 7279: gap of      100 bp
            *    7280     9066: contig of 1787 bp in length
            *    9067 9166: gap of      100 bp
            *    9167    11414: contig of 2248 bp in length
```

```
            *    11415 11514: gap of      100 bp
            *    11515    14617: contig of 3103 bp in length
            *    14618 14717: gap of      100 bp
            *    14718    18426: contig of 3709 bp in length
            *    18427 18526: gap of      100 bp
            *    18527    22011: contig of 3485 bp in length
            *    22012 22111: gap of      100 bp
            *    22112    26918: contig of 4807 bp in length
            *    26919 27018: gap of      100 bp
            *    27019    32120: contig of 5102 bp in length
            *    32121 32220: gap of      100 bp
            *    32221    37187: contig of 4967 bp in length
            *    37188 37287: gap of      100 bp
            *    37288    41477: contig of 4190 bp in length
            *    41478 41577: gap of      100 bp
            *    41578    46736: contig of 5159 bp in length
            *    46737 46836: gap of      100 bp
            *    46837    54164: contig of 7328 bp in length
            *    54165 54264: gap of      100 bp
            *    54265    58419: contig of 4155 bp in length
            *    58420 58519: gap of      100 bp
            *    58520    67538: contig of 9019 bp in length
            *    67539 67638: gap of      100 bp
            *    67639    74093: contig of 6455 bp in length
            *    74094 74193: gap of      100 bp
            *    74194    89734: contig of 15541 bp in length
            *    89735 89834: gap of      100 bp
            *    89835   109850: contig of 20016 bp in length
            *   109851 109950: gap of      100 bp
            *   109951   131927: contig of 21977 bp in length
            *   131928 132027: gap of      100 bp
            *   132028   163683: contig of 31656 bp in length.
FEATURES             Location/Qualifiers
     source          1. .163683
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1"
                     /clone="RP11-306I4"
                     /clone_lib="RPCI-11 Human Male BAC"
     misc_feature    1. .1333
                     /note="assembly_fragment"
     misc_feature    1434. .2643
                     /note="assembly_fragment"
     misc_feature    2744. .4634
                     /note="assembly_fragment"
     misc_feature    4735. .7179
                     /note="assembly_fragment"
     misc_feature    7280. .9066
                     /note="assembly_fragment"
     misc_feature    9167. .11414
                     /note="assembly_fragment"
     misc_feature    11515. .14617
                     /note="assembly_fragment"
     misc_feature    14718. .18426
                     /note="assembly_fragment"
     misc_feature    18527. .22011
                     /note="assembly_fragment"
     misc_feature    22112. .26918
                     /note="assembly_fragment"
     misc_feature    27019. .32120
                     /note="assembly_fragment"
     misc_feature    32221. .37187
                     /note="assembly_fragment"
     misc_feature    37288. .41477
                     /note="assembly_fragment"
     misc_feature    41578. .46736
                     /note="assembly_fragment"
     misc_feature    46837. .54164
                     /note="assembly_fragment"
     misc_feature    54265. .58419
                     /note="assembly_fragment
                     clone_end:SP6
                     vector_side:left"
     misc_feature    58520. .67538
                     /note="assembly_fragment"
     misc_feature    67639. .74093
                     /note="assembly_fragment
                     clone_end:T7
                     vector_side:right"
     misc_feature    74194. .89734
                     /note="assembly_fragment"
     misc_feature    89835. .109850
                     /note="assembly_fragment"
     misc_feature    109951. .131927
                     /note="assembly_fragment"
     misc_feature    132028. .163683
                     /note="assembly_fragment"
BASE COUNT    51900 a  32175 c  30125 g  47374 t   2109 others
ORIGIN


  Query Match             2.5%;  Score 229.2;  DB 64;  Length 163683;
  Best Local Similarity   76.2%;  Pred. No. 5.9e-28;
  Matches  282;  Conservative    0;  Mismatches   88;  Indels    0;  Gaps    0;

Qy   6858 tctgttgtttggttcagggattttatttttaattttatgagacagggtcttgctctgtcc 6917
          ||| || | |  |   |  ||| ||||| || ||||   ||||||||||||||||||| |
Db  93395 TCTTTTTTCTTTTATCGTTATTATATTTATATTTTTTCAAGACAGGGTCTTGCTCTGTTC 93336

Qy   6918 cccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggctca 6977
          |||||||||||||| ||||||  |||||||  || || ||||||||    |||||||| ||
Db  93335 CCCAGGCTGGAGTACAGTGGTGTGATCATGCCTTACCGCAGCCTCGATTTCCTGGGCACA 93276

Qy   6978 agcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgtgcc 7037
           | |||||||||||||  || | |||| ||||||||||||||||||||  |||||   |
Db  93275 GGAAATCCTCCTGCCTTGGCTTCCCAAGTAGCTGGGACTACAGGTGCGTACCACCACACG 93216

Qy   7038 tggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctggtct 7097
          ||||||| ||||   |||| |||||| || ||||| | ||  ||||||| | ||||||||
Db  93215 TGGCTAATTTTTATATTTTCTTGTAGAGATGGGGTTTTGTCATGTTGCCCAGGCTGGTCT 93156

Qy   7098 caaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgcaga 7157
          ||||||  ||| ||||||| |||| ||  ||||||||||||||||||||| ||||||||
Db  93155 CAAACTCCTGGGCTCAAGCCATCCGCCCACCTTGGCCTCCCAAAGTGCTGGGATTGCAGG 93096

Qy   7158 tgtgagccaccgtgcctggccagatttttctttattcttctttctttttttctttttgct 7217
          |||||| |||| ||||||| |   ||||| |||| ||   |||   |||   |||    |
Db  93095 TGTGAGGCACCATGCCTGGACTTTTTTTTTTTTTTTTTAACTTAAATTTAAATTTAAATT 93036

Qy   7218 ttcttgtctt 7227
          ||  | ||||
Db  93035 TTTATTTCTT 93026


RESULT  10
AC019269
LOCUS       AC019269   152511 bp    DNA             HTG       05-SEP-2000
DEFINITION  Homo sapiens clone RP11-8P13, WORKING DRAFT SEQUENCE, 10 unordered
            pieces.
ACCESSION   AC019269
VERSION     AC019269.4  GI:9966967
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 152511)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C. and Lander,E.
  TITLE     Homo sapiens, clone RP11-8P13
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 152511)
  AUTHORS   Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
            Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
            Boguslavkiy,L., Boukhgalter,B., Brown,A., Burkett,G., Castle,A.,
            Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
```

```
            DeArellano,K., Dewar,K., Domino,M., Doyle,M., Fenestor,J.,
            Ferreira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
            Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
            Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
            Landers,T., Lehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
            Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
            McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
            Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
            Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
            Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
            Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
            Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
            Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (31-DEC-1999) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT     On Sep 5, 2000 this sequence version replaced gi:7329377.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            -------------- Genome Center
                Center: Whitehead Institute/ MIT Center for Genome Research
                Center code: WIBR
                Web site: http://www-seq.wi.mit.edu
                Contact: sequence_submissions@genome.wi.mit.edu
            -------------- Project Information
                Center project name: L2991
                Center clone name: 8_P_13
            -------------- Summary Statistics
                Sequencing vector: M13; M77815; 96% of reads
                Sequencing vector: Plasmid; n/a; %-0.f%% of reads
                4.40579710144928Chemistry: Dye-terminator Big Dye; 100% of
            reads
                Assembly program: Phrap; version 0.960731
                Consensus quality: 148050 bases at least Q40
                Consensus quality: 150303 bases at least Q30
                Consensus quality: 151107 bases at least Q20
                Insert size: 163000; agarose-fp
                Insert size: 151611; sum-of-contigs
                Quality coverage: 5.5 in Q20 bases; agarose-fp
                Quality cove.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 10 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     8618: contig of 8618 bp in length
            *     8619 8718: gap of      100 bp
            *     8719    13037: contig of 4319 bp in length
            *    13038 13137: gap of      100 bp
            *    13138    19064: contig of 5927 bp in length
            *    19065 19164: gap of      100 bp
            *    19165    27544: contig of 8380 bp in length
            *    27545 27644: gap of      100 bp
            *    27645    36519: contig of 8875 bp in length
            *    36520 36619: gap of      100 bp
            *    36620    43866: contig of 7247 bp in length
            *    43867 43966: gap of      100 bp
            *    43967    59245: contig of 15279 bp in length
            *    59246 59345: gap of      100 bp
            *    59346    81568: contig of 22223 bp in length
            *    81569 81668: gap of      100 bp
            *    81669   110160: contig of 28492 bp in length
            *   110161 110260: gap of      100 bp
            *   110261   152511: contig of 42251 bp in length.
FEATURES            Location/Qualifiers
     source         1. .152511
                    /organism="Homo sapiens"
                    /db_xref="taxon:9606"
                    /clone="RP11-8P13"
                    /clone_lib="RPCI-11 Human Male BAC"
     misc_feature   1. .8618
                    /note="assembly_fragment
                    clone_end:SP6
                    vector_side:left"
     misc_feature   8719. .13037
                    /note="assembly_fragment"
     misc_feature   13138. .19064
                    /note="assembly_fragment"
     misc_feature   19165. .27544
                    /note="assembly_fragment"
     misc_feature   27645. .36519
                    /note="assembly_fragment"
     misc_feature   36620. .43866
                    /note="assembly_fragment"
     misc_feature   43967. .59245
                    /note="assembly_fragment"
     misc_feature   59346. .81568
                    /note="assembly_fragment"
     misc_feature   81669. .110160
                    /note="assembly_fragment"
     misc_feature   110261. .152511
                    /note="assembly_fragment"
BASE COUNT    45636 a  30476 c  30973 g  44525 t    901 others
ORIGIN


  Query Match             2.5%;  Score 226.2;  DB 60;  Length 152511;
  Best Local Similarity   80.7%;  Pred. No. 1.8e-27;
  Matches  264;  Conservative    0;  Mismatches   63;  Indels    0;  Gaps    0;

Qy    6866 ttggttcagggatttttatttttaattttatgagacagggtcttgctctgtcccccaggct 6925
           ||  || |   ||||  |||||  |||| ||||||||| |||| ||||||||| ||||||||
Db 119560 TTAATTTACTCATTTCTTTTTTTCTTTTTTGAGACAGAGTCTCGCTCTGTCGCCCAGGCT 119619

Qy    6926 ggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggctcaagcaatcc 6985
           |||||||||||||||||| | |||||||||| ||||   |||||||| ||||||||| |
Db 119620 GGAGTGCAGTGGCATGATCTTGGCTCACTGCAACCTCTGCCTCCTGGGTTCAAGCAATTC 119679

Qy    6986 tcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgtgcctggctaac 7045
           |||||||||| || |  | ||||||||||||||| || | ||||| ||||||||||
Db 119680 TCCTGCCTCAGTCTCCTGAGTAGCTGGGACTACAGGCGCCCACCACCATGCCTGGCTAAT 119739

Qy    7046 ttttcatttttttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttg 7105
           ||||  | ||||| |||| |||||||| ||    ||||  | | ||||||||  ||||
Db 119740 TTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTAGCCAGGCTGGTCTTGAACTCC 119799

Qy    7106 tggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgcagatgtgagcc 7165
           || |||||||  |||||||||||| |||||||||||||||| |||| |||  |||||||
Db 119800 TGACCTCAAGTGATCCACCTGCCTCGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCC 119859

Qy    7166 accgtgcctggccagatttttctttta 7192
           |||||||||||||   ||  ||||||
Db 119860 ACCGTGCCTGGCCATTATTAACTTTTA 119886


RESULT  11
AC024740
LOCUS       AC024740   179538 bp    DNA             HTG       07-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone RP11-801D6, WORKING DRAFT
            SEQUENCE, 17 unordered pieces.
ACCESSION   AC024740
VERSION     AC024740.3  GI:8570405
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 179538)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 179538)
```

```
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (01-MAR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Jun 17, 2000 this sequence version replaced gi:7263917.

            ---------- Genome Center ----------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            ---------- Project Information ----------
            Center project name: H_NH0801D06
            ---------- Summary Statistics  ----------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 97% of reads
            Chemistry: Dye-terminator Big Dye; 3% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 173830 bases at least Q40
            Consensus quality: 175557 bases at least Q30
            Consensus quality: 176446 bases at least Q20
            Insert size: 242000; agarose-fp
            Insert size: 177938; sum-of-contigs
            Quality coverage: 5.32 in Q20 bases; agarose-fp
            Quality coverage: 6.01 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     1801: contig of 1801 bp in length
            *     1802     1901: gap of unknown length
            *     1902     4641: contig of 2740 bp in length
            *     4642     4741: gap of unknown length
            *     4742     9679: contig of 4938 bp in length
            *     9680     9779: gap of unknown length
            *     9780    13392: contig of 3613 bp in length
            *    13393    13492: gap of unknown length
            *    13493    17164: contig of 3672 bp in length
            *    17165    17264: gap of unknown length
            *    17265    21510: contig of 4246 bp in length
            *    21511    21610: gap of unknown length
            *    21611    27089: contig of 5479 bp in length
            *    27090    27189: gap of unknown length
            *    27190    35636: contig of 8447 bp in length
            *    35637    35736: gap of unknown length
            *    35737    45537: contig of 9801 bp in length
            *    45538    45637: gap of unknown length
            *    45638    56451: contig of 10814 bp in length
            *    56452    56551: gap of unknown length
            *    56552    66802: contig of 10251 bp in length
            *    66803    66902: gap of unknown length
            *    66903    79118: contig of 12216 bp in length
            *    79119    79218: gap of unknown length
            *    79219    94355: contig of 15137 bp in length
            *    94356    94455: gap of unknown length
            *    94456   111274: contig of 16819 bp in length
            *   111275   111374: gap of unknown length
            *   111375   130600: contig of 19226 bp in length
            *   130601   130700: gap of unknown length
            *   130701   150278: contig of 19578 bp in length
            *   150279   150378: gap of unknown length
            *   150379   179538: contig of 29160 bp in length.
FEATURES             Location/Qualifiers
     source          1. .179538
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="RP11-801D6"
     misc_feature    1. .1801
                     /note="assembly_name:Contig13"
     misc_feature    1902. .4641
                     /note="assembly_name:Contig14"
     misc_feature    4742. .9679
                     /note="assembly_name:Contig15"
     misc_feature    9780. .13392
                     /note="assembly_name:Contig16"
     misc_feature    13493. .17164
                     /note="assembly_name:Contig17"
     misc_feature    17265. .21510
                     /note="assembly_name:Contig18"
     misc_feature    21611. .27089
                     /note="assembly_name:Contig19"
     misc_feature    27190. .35636
                     /note="assembly_name:Contig20"
     misc_feature    35737. .45537
                     /note="assembly_name:Contig21"
     misc_feature    45638. .56451
                     /note="assembly_name:Contig22"
     misc_feature    56552. .66802
                     /note="assembly_name:Contig23"
     misc_feature    66903. .79118
                     /note="assembly_name:Contig24"
     misc_feature    79219. .94355
                     /note="assembly_name:Contig25"
     misc_feature    94456. .111274
                     /note="assembly_name:Contig26"
     misc_feature    111375. .130600
                     /note="assembly_name:Contig27"
     misc_feature    130701. .150278
                     /note="assembly_name:Contig28
                     clone_end:SP6
                     vector_side:left"
     misc_feature    150379. .179538
                     /note="assembly_name:Contig29"
BASE COUNT    42230 a  47023 c  46289 g  42392 t   1604 others
ORIGIN


  Query Match             2.5%;  Score 225.6;  DB 63;  Length 179538;
  Best Local Similarity   67.4%;  Pred. No. 2.2e-27;
  Matches  318;  Conservative    0;  Mismatches  154;  Indels    0;  Gaps    0;

Qy    6853 gtttctctgttgtttggttcagggattttatttttaattttatgagacagggtcttgctc 6912
           ||||   | |  |||||||      |||  ||| |  |||| ||||||||| ||||  |||
Db  167103 GTTTTGTTTTGTTTTGGTTTGTTTTTTTGTTTTGTTGTTTTTTGAGACAGAGTCTCACTC 167162

Qy    6913 tgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgg 6972
           ||||  ||||||||||||||||||||    ||||   |||||||||| ||||   |||||||
Db  167163 TGTCACCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCCGCCTCCTGG 167222

Qy    6973 gctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccacc 7032
           | ||||   || ||||||||||||||||    | |||||||||| |||||| || |||||||
Db  167223 GTTCAACGGATTCTCCTGCCTCAGCCTCTGGAGTAGCTGGGATTACAGGCGCACGCCACC 167282

Qy    7033 gtgcctggctaacttttcatttttttttgtagggacggggtctcgttttgttgccaaagct 7092
             ||||||||||  ||||   | |||||  |||| || |||||  |||    |||||  | | |||
Db  167283 ACGCCTGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCT 167342

Qy    7093 ggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagatt 7152
           ||||||||||  ||||||||  |   |||| |||  |||  ||||||||||||||| ||||
Db  167343 GGTCTCAAACTCCTGGCCTCAGGTGATCCGCCTACCTCAGCCTCCCAAAGTGCTGGGATT 167402

Qy    7153 gcagatgtgagccaccgtgcctggccagatttttctttttattcttctttctttttctttt 7212
            |||   |||||||||| | ||| ||||   |||| |||  | | |||||  |||   ||
Db  167403 ACAGGCATGAGCCACCGCGTCTGTCCAGTCCTTTCCTTTCCTTTCCTTTCCATTTTCCTT 167462

Qy    7213 ttgctttcttgtcttttcagaagcaagccagacctagcaggctgttccatgttctatttt 7272
           |   || |||  |||||      |    |      | |   ||  |  |   |||| | ||
Db  167463 TCCTTTCCTTTCCTTTTTCCTTTCCTTTCCTTTTTTTCTCTCTCTCTCTCTTTCTTTCTT 167522
```

```
Qy   7273 tgactgtagccacagctgctgttctcaggacagcatcccttcccacatgcct 7324
          |  || |  |     ||     || |    |  |   | ||||| | | |||
Db 167523 TCTCTTTCTCTCTTTCTTTCTCTCCCTTCCCTTCCCTCTTTCCCTCTTTCCT 167574

RESULT 12
AL355392/c                  /
LOCUS       AL355392   159440 bp    DNA             PRI       13-SEP-2000
DEFINITION  Human DNA sequence from clone RP5-1187J4 on chromosome
            20q11.1-11.23, complete sequence.
ACCESSION   AL355392 AL117331
VERSION     AL355392.7  GI:10178502
KEYWORDS    HTG.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 159440)
  AUTHORS   Bird,C.
  TITLE     Direct Submission
  JOURNAL   Submitted (12-SEP-2000) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
            requests: clonerequest@sanger.ac.uk
COMMENT     On Sep 15, 2000 this sequence version replaced gi:9213549.
            During sequence assembly data is compared from overlapping clones.
            Where differences are found these are annotated as variations
            together with a note of the overlapping clone name. Note that the
            variation annotation may not be found in the sequence submission
            corresponding to the overlapping clone, as we submit sequences with
            only a small overlap as described above.
            This sequence has been finished according to sequence map criteria
            as follows. An attempt is made to resolve all sequencing problems,
            such as compressions and repeats, but not necessarily within known
            annotated human repeat sequence elements (e.g. Alu). Where the
            sequence is ambiguous, there is an annotation using the 'unsure'
            feature key.
            The following abbreviations are used to associate primary accession
            numbers given in the feature table with their source databases:
            Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
            on the WORMPEP database can be found at
            http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
            was generated from part of bacterial clone contigs of human
            chromosome 20, constructed by the Sanger Centre Chromosome 20
            Mapping Group.  Further information can be found at
            http://www.sanger.ac.uk/HGP/Chr20
            RP5-1187J4 is from the library RPCI-5 constructed at the Roswell
            Park Cancer Institute by the group of Pieter de Jong. For further
            details see http://bacpac.med.buffalo.edu/
            VECTOR: pCYPAC2
            This sequence is the entire insert of clone RP5-1187J4 The true
            left end of clone RP11-327D19 is at 62286 in this sequence.
FEATURES             Location/Qualifiers
     source          1. .159440
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="20"
                     /map="q11.1-11.23"
                     /clone="RP5-1187J4"
                     /clone_lib="RPCI-5"
     repeat_region   816. .1129
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   1207. .1534
                     /note="AluSx repeat: matches 1. .311 of consensus"
     repeat_region   1535. .1848
                     /note="157 copies 2 mer ga 61% conserved"
     misc_feature    1757. .1824
                     /note="Tandem repeat. weak data"
     repeat_region   1917. .2031
                     /note="MER5B repeat: matches 64. .175 of consensus"
     repeat_region   2415. .2622
                     /note="MER20 repeat: matches 1. .218 of consensus"
     repeat_region   2811. .3103
                     /note="AluSx repeat: matches 1. .296 of consensus"
     repeat_region   3703. .3908
                     /note="MIR repeat: matches 47. .250 of consensus"
     misc_feature    3817. .3915
                     /note="match: STS: Em:G10242"
     misc_feature    complement(join(3917. .4006,31119. .31326))
                     /note="match: STS: Em:U75536"
     misc_feature    3917. .4143
                     /note="match: STS: Em:G10242"
     repeat_region   4399. .4569
                     /note="MIR repeat: matches 65. .242 of consensus"
     repeat_region   4662. .4691
                     /note="15 copies 2 mer tt 86% conserved"
     repeat_region   4913. .4977
                     /note="U6 repeat: matches 1. .65 of consensus"
     repeat_region   5424. .5616
                     /note="MIR repeat: matches 28. .251 of consensus"
     repeat_region   6034. .6144
                     /note="AluSg/x repeat: matches 186. .302 of consensus"
     repeat_region   6200. .6308
                     /note="MIR repeat: matches 47. .155 of consensus"
     repeat_region   6467. .6770
                     /note="AluJb repeat: matches 1. .303 of consensus"
     misc_feature    complement(6761. .7253)
                     /note="match: GSS: Em:AQ149447"
     repeat_region   8416. .8769
                     /note="L2 repeat: matches 2217. .2616 of consensus"
     repeat_region   9663. .9761
                     /note="MIR repeat: matches 50. .165 of consensus"
     repeat_region   10562. .10676
                     /note="L2 repeat: matches 2620. .2750 of consensus"
     repeat_region   10970. .11009
                     /note="20 copies 2 mer ac 85% conserved"
     repeat_region   11192. .11497
                     /note="AluY repeat: matches 3. .305 of consensus"
     misc_feature    complement(11466. .11766)
                     /note="match: GSS: Em:AQ060606"
     misc_feature    complement(11631. .11769)
                     /note="match: GSS: Em:AQ427662"
     misc_feature    11766. .12192
                     /note="match: GSS: Em:AQ403818"
     misc_feature    11768. .12108
                     /note="match: GSS: Em:AQ373301"
     misc_feature    11775. .12108
                     /note="match: GSS: Em:AQ662246"
     repeat_region   12252. .12360
                     /note="MIR repeat: matches 60. .170 of consensus"
     repeat_region   12970. .13022
                     /note="MIR repeat: matches 85. .138 of consensus"
     repeat_region   13116. .13474
                     /note="L2 repeat: matches 2386. .2750 of consensus"
     repeat_region   13475. .13514
                     /note="MSTA repeat: matches 388. .425 of consensus"
     repeat_region   13515. .13826
                     /note="AluSq repeat: matches 1. .307 of consensus"
     repeat_region   13827. .14198
                     /note="MSTA repeat: matches 1. .388 of consensus"
     repeat_region   14199. .14264
                     /note="L2 repeat: matches 2322. .2386 of consensus"
     repeat_region   14322. .14833
                     /note="256 copies 2 mer ca 68% conserved"
     repeat_region   14836. .14871
                     /note="18 copies 2 mer ac 88% conserved"
     repeat_region   14896. .14995
                     /note="L2 repeat: matches 2607. .2708 of consensus"
     misc_feature    complement(15031. .15459)
                     /note="match: GSS: Em:AQ694930"
     repeat_region   15294. .15407
                     /note="MIR repeat: matches 48. .160 of consensus"
     repeat_region   16063. .16274
                     /note="MIR repeat: matches 20. .253 of consensus"
     repeat_region   17156. .17282
                     /note="LTR16C repeat: matches 113. .236 of consensus"
     repeat_region   17350. .17524
```

```
                     /note="FRAM repeat: matches 1. .175 of consensus"
     repeat_region   17527. .17670
                     /note="LTR16C repeat: matches 242. .387 of consensus"
     misc_feature    complement(17559. .18000)
                     /note="match: GSS: Em:AQ349609"
     misc_feature    complement(17617. .18004)
                     /note="match: GSS: Em:AQ350711"
     misc_feature    18109. .18565
                     /note="match: STS: Em:HS812D7T"
     repeat_region   18262. .18574
                     /note="AluSx repeat: matches 1. .312 of consensus"
     repeat_region   18660. .18721
                     /note="MIR repeat: matches 92. .147 of consensus"
     repeat_region   19233. .19535
                     /note="AluSg repeat: matches 1. .303 of consensus"
     repeat_region   19911. .20016
                     /note="MIR repeat: matches 47. .152 of consensus"
     repeat_region   20023. .20252
                     /note="MIR repeat: matches 5. .249 of consensus"
     misc_feature    complement(20661. .21112)
                     /note="match: GSS: Em:AQ019396"
     misc_feature    complement(20698. .21115)
                     /note="match: GSS: Em:AQ178024"
     repeat_region   20940. .21166
                     /note="MIR repeat: matches 25. .249 of consensus"
     misc_feature    21105. .21680
                     /note="match: STS: Em:HS156A22T"
     repeat_region   22012. .22288
                     /note="AluJb repeat: matches 13. .290 of consensus"
     repeat_region   22356. .22515
                     /note="MIR repeat: matches 48. .221 of consensus"
     repeat_region   22508. .22708
                     /note="L2 repeat: matches 2223. .2435 of consensus"
     misc_feature    complement(22577. .22968)
                     /note="match: GSS: Em:AQ508829"
     misc_feature    22977. .23573
                     /note="match: GSS: Em:AQ281488"
     repeat_region   22998. .23126
                     /note="L2 repeat: matches 2157. .2280 of consensus"
     repeat_region   23152. .23316
                     /note="L2 repeat: matches 2371. .2537 of consensus"
     repeat_region   23399. .23864
                     /note="L2 repeat: matches 2250. .2750 of consensus"
     repeat_region   24480. .24558
                     /note="MLT1B repeat: matches 318. .390 of consensus"
     repeat_region   24560. .24848
                     /note="AluJb repeat: matches 4. .291 of consensus"
     repeat_region   24868. .25158
                     /note="AluSx repeat: matches 14. .304 of consensus"
     repeat_region   25171. .25326
                     /note="FAM repeat: matches 5. .174 of consensus"
     repeat_region   25621. .25684
                     /note="MIR repeat: matches 94. .157 of consensus"
     repeat_region   26144. .26233
                     /note="MIR repeat: matches 50. .140 of consensus"
     repeat_region   26337. .26421
                     /note="MIR repeat: matches 66. .139 of consensus"
     repeat_region   26422. .26719
                     /note="AluSx repeat: matches 1. .298 of consensus"
     repeat_region   26720. .26763
                     /note="MIR repeat: matches 139. .185 of consensus"

  Query Match             2.5%;  Score 224.6;  DB 34;  Length 159440;
  Best Local Similarity   77.5%;  Pred. No. 3.2e-27;
  Matches  272;  Conservative    0;  Mismatches   79;  Indels    0;  Gaps    0;

Qy   6878 ttttatttttaatttttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
           |||| |||||  ||||  ||||||||||||||||||||| ||||||||||||||||||||
Db 127857 TTTTTTTTTTTTTTTTTGGAGACAGGGTCTTGCTCTGTCGCCCAGGCTGGAGTGCAGTGG 127798

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
           ||| ||| | ||| ||||||| ||||   |||||||||||||||||||||||  |||||||
Db 127797 CATAATCTTGGCTTACTGCAACCTCTGCCTCCTGGGCTCAAGCAATCCTCCCACCTCAGC 127738

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || || | ||||||||||||||||| || | | ||| |||||||||||  |||  | ||||
Db 127737 CTCCCGAGTAGCTGGGACTACAGGCGCACACTACCATGCCTGGCTAATGTTTTGTGTTTT 127678

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| |||  ||| | | | ||||||| || |||||||  ||||  ||| |||||||||
Db 127677 TGGTAGAGACAAGGTTTTGCTATGTTGCCCAACCTGGTCTTGAACTCCTGGGCTCAAGCA 127618

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
           ||  ||||||| || |||||||||||||| |||| |||  |||||||||| || |||||
Db 127617 GTCTGCCTGCCTCGGGCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCATGTCTGGC 127558

Qy   7178 cagatttttctttttattcttctttctttttctttttgctttcttgtcttt 7228
             | || |   |  ||  | ||| || || | || |  ||| || |  ||
Db 127557 TGTAGTTATTAATCTTTTATTTTTATTATTTTATTATTTTTTATTTTTATT 127507


RESULT  13
AC016695
LOCUS       AC016695   177560 bp    DNA             HTG       05-OCT-2000
DEFINITION  Homo sapiens chromosome 7 clone RP11-126N23, WORKING DRAFT
            SEQUENCE, 14 unordered pieces.
ACCESSION   AC016695
VERSION     AC016695.8  GI:10645615
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 177560)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 177560)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-DEC-1999) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Oct 5, 2000 this sequence version replaced gi:9800603.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0126N23
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 72%
            Sequencing vector: plasmid; 7%
            Chemistry: Dye-primer ET; 72% of reads
            Chemistry: Dye-terminator Big Dye; 7% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 168896 bases at least Q40
            Consensus quality: 171731 bases at least Q30
            Consensus quality: 173249 bases at least Q20
            Insert size: 186Kb; agarose-fp
            Insert size: 176260; sum-of-contigs
            Quality coverage: 5.15 in Q20 bases; agarose-fp
            Quality coverage: 5.49 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 14 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *       1     2804: contig of 2804 bp in length
            *    2805     2904: gap of unknown length
            *    2905     5617: contig of 2713 bp in length
```

```
            *     5618     5717: gap of unknown length
            *     5718     9655: contig of 3938 bp in length
            *     9656     9755: gap of unknown length
            *     9756    13384: contig of 3629 bp in length
            *    13385    13484: gap of unknown length
            *    13485    18533: contig of 5049 bp in length
            *    18534    18633: gap of unknown length
            *    18634    25742: contig of 7109 bp in length
            *    25743    25842: gap of unknown length
            *    25843    34853: contig of 9011 bp in length
            *    34854    34953: gap of unknown length
            *    34954    46300: contig of 11347 bp in length
            *    46301    46400: gap of unknown length
            *    46401    58091: contig of 11691 bp in length
            *    58092    58191: gap of unknown length
            *    58192    78388: contig of 20197 bp in length
            *    78389    78488: gap of unknown length
            *    78489   101520: contig of 23032 bp in length
            *   101521   101620: gap of unknown length
            *   101621   123033: contig of 21413 bp in length
            *   123034   123133: gap of unknown length
            *   123134   147534: contig of 24401 bp in length
            *   147535   147634: gap of unknown length
            *   147635   177560: contig of 29926 bp in length.
FEATURES             Location/Qualifiers
     source          1. .177560
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="7"
                     /clone="RP11-126N23"
     misc_feature    1. .2804
                     /note="assembly_name:Contig19"
     misc_feature    2905. .5617
                     /note="assembly_name:Contig20"
     misc_feature    5718. .9655
                     /note="assembly_name:Contig21"
     misc_feature    9756. .13384
                     /note="assembly_name:Contig22"
     misc_feature    13485. .18533
                     /note="assembly_name:Contig23"
     misc_feature    18634. .25742
                     /note="assembly_name:Contig24"
     misc_feature    25843. .34853
                     /note="assembly_name:Contig25"
     misc_feature    34954. .46300
                     /note="assembly_name:Contig26
                     clone_end:SP6
                     vector_side:right"
     misc_feature    46401. .58091
                     /note="assembly_name:Contig27"
     misc_feature    58192. .78388
                     /note="assembly_name:Contig28"
     misc_feature    78489. .101520
                     /note="assembly_name:Contig29"
     misc_feature    101621. .123033
                     /note="assembly_name:Contig30"
     misc_feature    123134. .147534
                     /note="assembly_name:Contig31"
     misc_feature    147635. .177560
                     /note="assembly_name:Contig32"
BASE COUNT    47358 a  39766 c  39766 g  49364 t   1306 others
ORIGIN


  Query Match             2.5%;  Score 224.2;  DB 59;  Length 177560;
  Best Local Similarity   76.8%;  Pred. No. 3.6e-27;
  Matches  274;  Conservative    0;  Mismatches   83;  Indels    0;  Gaps    0;

Qy   6863 tgtttggttcagggatttttatttttaattttatgagacagggtcttgctctgtccccccag 6922
          |||||  |||     |||||  |||||  ||||  |||||||||||||| | ||||||  |||||
Db  83262 TGTTTTGTTTTTTAATTTTTTTTTTTTTTTTTTGAGACAGGGTCTCGTTCTGTCACCCAG 83321

Qy  6923 gctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggctcaagcaa 6982
```

```
          |||||| |||||||||||| |||| |||||||||||||| |   ||||||| |||||||||
Db  83322 GCTGGAATGCAGTGGCATAATCACAGCTCACTGCAGCCACCACCTCCTGGACTCAAGCAA 83381

Qy   6983 tcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgtgcctggct 7042
          ||||||  |||||||||| |||| ||||||||||||||||  || | ||||   ||  |||
Db  83382 TCCTCCCACCTCAGCCTCCCAAGTAGCTGGGACTACAGGCACGTGTCACCACACCCAGCT 83441

Qy   7043 aacttttcatttttttttgtagggacggggtctcgttttgttgccaaagctggtctcaaac 7102
          || ||||   |||||||||||| || ||||| ||    ||||||| | |||||||| |||
Db  83442 AATTTTTGTATTTTTTTGTAGAGATGGGGTTTCACCATGTTGCCCAGGCTGGTCTTGAAC 83501

Qy   7103 ttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgcagatgtga 7162
          |  ||| ||||||||||   |||||||||||||| ||||| |||| ||||| |||   |||
Db  83502 TCCTGGGCTCAAGCAATATGCCTGCCTTGGCCTTCCAAACTGCTCAGATTACAGGCATGA 83561

Qy   7163 gccaccgtgcctggccagatttttcttttattcttctttctttttttcttttttgcttt 7219
          ||||||| ||||||||||  |     | |  | | | || |||   ||||||  | |
Db  83562 GCCACCGCGCCTGGCCAGTCTACGTATCTTATATGCATTATTTCATTTTTTTAATGT 83618


RESULT  14
AC026803
LOCUS       AC026803   227949 bp    DNA             HTG       18-JUL-2000
DEFINITION  Homo sapiens chromosome 19 clone CTD-2639E6, WORKING DRAFT
            SEQUENCE, 13 ordered pieces.
ACCESSION   AC026803
VERSION     AC026803.4  GI:9256691
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 227949)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Sequencing of Human Chromosome 19
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 227949)
  AUTHORS   DOE Joint Genome Institute.
  TITLE     Direct Submission
  JOURNAL   Submitted (24-MAR-2000) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
COMMENT     On Jul 18, 2000 this sequence version replaced gi:8576168.
            -------Genome Center
            Center: Joint Genome Institute
            Center Code: JGI
            Web site: http://www.jgi.doe.gov
            -------
            Project Information
            Center Project Name: 843459, BC808483
            Center clone name: CITB-E1_2639E6
            -------
            Summary Statistics
            Consensus quality: 217198 bases at least Q40
            Consensus quality: 224116 bases at least Q30
            Consensus quality: 225875 bases at least Q20
            Estimated insert size: 226670; agarose-fp estimation
            Estimated insert size: 227399; sum-of-contigs estimation
            Quality coverage:  6.85 in Q20 bases; agarose-fp estimation
            Quality coverage:  6.83 in Q20 bases; sum-of-contigs estimation.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 13 contigs. Gaps between the contigs
            * are represented as runs of N. The order of the pieces
            * is believed to be correct as given, however the sizes
            * of the gaps between them are based on estimates that have
            * provided by the submittor.
            * This sequence will be replaced
            * by the finished sequence as soon as it is available and
            * the accession number will be preserved.
            *        1    21840: contig of 21840 bp in length
            *    21841    21940: gap of unknown length
            *    21941    31559: contig of 9619 bp in length
            *    31560    31659: gap of unknown length
            *    31660    36187: contig of 4528 bp in length
```

```
            *    36188    36287: gap of unknown length
            *    36288    59202: contig of 22915 bp in length
            *    59203    59302: gap of unknown length
            *    59303    69077: contig of 9775 bp in length
            *    69078    69177: gap of unknown length
            *    69178    80927: contig of 11750 bp in length
            *    80928    81027: gap of unknown length
            *    81028   138192: contig of 57165 bp in length
            *   138193   138292: gap of unknown length
            *   138293   160224: contig of 21932 bp in length
            *   160225   160324: gap of unknown length
            *   160325   161182: contig of 858 bp in length
            *   161183   161282: gap of unknown length
            *   161283   162861: contig of 1579 bp in length
            *   162862   162961: gap of unknown length
            *   162962   221619: contig of 58658 bp in length
            *   221620   221719: gap of unknown length
            *   221720   225761: contig of 4042 bp in length
            *   225762   225861: gap of unknown length
            *   225862   227949: contig of 2088 bp in length.
FEATURES             Location/Qualifiers
     source          1. .227949
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="19"
                     /clone="CTD-2639E6"
                     /clone_lib="CalTech human BAC library D"
BASE COUNT    54671 a  58658 c  58905 g  54511 t   1204 others
ORIGIN


  Query Match             2.5%;  Score 224;  DB 65;  Length 227949;
  Best Local Similarity   67.2%;  Pred. No. 3.6e-27;
  Matches  317;  Conservative    0;  Mismatches  155;  Indels    0;  Gaps    0;

Qy   6853 gtttctctgttgtttggttcagggatttattttttaattttatgagacagggtcttgctc 6912
          ||||   | |  |||||||      |||  ||| |  |||| ||||||||| ||||  |||
Db 123573 GTTTTGTTTTGTTTTGGTTTGTTTTTTTGTTTTGTTGTTTTTTGAGACAGAGTCTCTCTC 123632

Qy   6913 tgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgg 6972
          |||| ||||||||||||||||||    ||||   |||||||||| ||||   |||||||
Db 123633 TGTCACCCAGGCTGGAGTGCAGTGGTGCGATCTCGGCTCACTGCAACCTCCGCCTCCTGG 123692

Qy   6973 gctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccacc 7032
          | ||||   || ||||||||||||||    | |||||||| |||||| || |||||||
Db 123693 GTTCAACGGATTCTCCTGCCTCAGCCTCTGGAGTAGCTGGGATTACAGGCGCACGCCACC 123752

Qy   7033 gtgcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagct 7092
          ||||||||| ||||  | ||||| |||| || ||||| |||   |||||  | | |||
Db 123753 ACGCCTGGCTAATTTTTTGTATTTTTAGTAGAGATGGGGTTTCGCCATGTTGGCCAGGCT 123812

Qy   7093 ggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagatt 7152
          |||||||||| |||||||| |  |||| ||| |||  |||||||||||||||| ||||
Db 123813 GGTCTCAAACTCCTGGCCTCAGGTGATCCGCCTACCTCAGCCTCCCAAAGTGCTGGGATT 123872

Qy   7153 gcagatgtgagccaccgtgcctggccagatttttcttttattcttctttcttttttctttt 7212
           |||   |||||||||| | ||| ||||   |||| |||  | | |||||  |||   ||
Db 123873 ACAGGCATGAGCCACCGCGTCTGTCCAGTCCTTTCCTTTCCTTTCCTTTCCATTTTCCTT 123932

Qy   7213 ttgctttcttgtctttttcagaagcaagccagacctagcaggctgttccatgttctatttt 7272
          |  || |||  |||||      |    |      |     || |  |  |||| | ||
Db 123933 TCCTTTCCTTTCCTTTTTCCTTTCCTTTCCTTTTTTTTTCTCTCTCTCTCTTTCTTTCTT 123992

Qy   7273 tgactgtagccacagctgctgttctcaggacagcatcccttcccacatgcct 7324
          |  || |  |     ||     || |    |  |   | ||||| | | |||
Db 123993 TCTCTTTCTCTCTTTCTTTCTCTCCCTTCCCTTCCCTCTTTCCCTCTTTCCT 124044


RESULT  15
AC036235
LOCUS       AC036235   160659 bp    DNA             HTG       07-JUL-2000
DEFINITION  Homo sapiens chromosome 6 clone RP11-113D12, WORKING DRAFT
            SEQUENCE, 11 unordered pieces.

ACCESSION   AC036235
VERSION     AC036235.2  GI:7574887
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 160659)
  AUTHORS   Waterston,R.H.
  TITLE     The sequence of Homo sapiens clone
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 160659)
  AUTHORS   Waterston,R.H.
  TITLE     Direct Submission
  JOURNAL   Submitted (07-APR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT     On Apr 15, 2000 this sequence version replaced gi:7523924.

            ----------------- Genome Center -----------------
            Center: Washington University Genome Sequencing Center
            Center code: WUGSC
            Web site:http://genome.wustl.edu/gsc/index.shtml
            -------------- Project Information --------------
            Center project name: H_NH0113D12
            -------------- Summary Statistics  --------------
            Sequencing vector: M13; 100%
            Sequencing vector: plasmid; 0%
            Chemistry: Dye-primer ET; 100% of reads
            Chemistry: Dye-terminator Big Dye; 0% of reads
            Assembly program: Phrap; version 0.990319
            Consensus quality: 156067 bases at least Q40
            Consensus quality: 157279 bases at least Q30
            Consensus quality: 158052 bases at least Q20
            Insert size: 168000; agarose-fp
            Insert size: 159659; sum-of-contigs
            Quality coverage: 4.95 in Q20 bases; agarose-fp
            Quality coverage: 5.24 in Q20 bases; sum-of-contigs
            ------------------------------------------------.
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 11 contigs. The true order of the pieces
            * is not known and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence
            * as soon as it is available and the accession number will
            * be preserved.
            *        1     2033: contig of 2033 bp in length
            *     2034     2133: gap of unknown length
            *     2134     6279: contig of 4146 bp in length
            *     6280     6379: gap of unknown length
            *     6380    11045: contig of 4666 bp in length
            *    11046    11145: gap of unknown length
            *    11146    19290: contig of 8145 bp in length
            *    19291    19390: gap of unknown length
            *    19391    31665: contig of 12275 bp in length
            *    31666    31765: gap of unknown length
            *    31766    47421: contig of 15656 bp in length
            *    47422    47521: gap of unknown length
            *    47522    62750: contig of 15229 bp in length
            *    62751    62850: gap of unknown length
            *    62851    79257: contig of 16407 bp in length
            *    79258    79357: gap of unknown length
            *    79358    97520: contig of 18163 bp in length
            *    97521    97620: gap of unknown length
            *    97621   123364: contig of 25744 bp in length
            *   123365   123464: gap of unknown length
            *   123465   160659: contig of 37195 bp in length.
FEATURES             Location/Qualifiers
     source          1. .160659
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="6"
```

```
                     /clone="RP11-113D12"
     misc_feature    1. .2033
                     /note="assembly_name:Contig1"
     misc_feature    2134. .6279
                     /note="assembly_name:Contig2"
     misc_feature    6380. .11045
                     /note="assembly_name:Contig3
                     clone_end:T7
                     vector_side:left"
     misc_feature    11146. .19290
                     /note="assembly_name:Contig4"
     misc_feature    19391. .31665
                     /note="assembly_name:Contig5"
     misc_feature    31766. .47421
                     /note="assembly_name:Contig6"
     misc_feature    47522. .62750
                     /note="assembly_name:Contig7"
     misc_feature    62851. .79257
                     /note="assembly_name:Contig8"
     misc_feature    79358. .97520
                     /note="assembly_name:Contig9"
     misc_feature    97621. .123364
                     /note="assembly_name:Contig10"
     misc_feature    123465. .160659
                     /note="assembly_name:Contig11
                     clone_end:SP6
                     vector_side:right"
BASE COUNT    49813 a  30389 c  32438 g  47016 t   1003 others
ORIGIN


  Query Match             2.5%;  Score 223.6;  DB 66;  Length 160659;
  Best Local Similarity   74.1%;  Pred. No. 4.6e-27;
  Matches  283;  Conservative    0;  Mismatches   99;  Indels    0;  Gaps    0;

Qy   6842 taaaaggtggagtttctctgttgtttggttcagggatttttatttttaattttatgagaca 6901
          ||||||   |  ||| ||| || |||  || |     |||  ||||| |||   |||||.|
Db  46115 TAAAAGACTGTATTTTTCTTTTCTTTTCTTTATTTTATTTTATTTTATTTTAGGGAGACA 46174

Qy   6902 gggtcttgctctgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcct 6961
          | ||| |||||| |||||||||||||||||||||||||| |||| | |||||||||| |||
Db  46175 GAGTCCTGCTCTATCCCCCAGGCTGGAGTGCAGTGGCACGATCTTGGCTCACTGCAACCT 46234

Qy   6962 catactcctgggctcaagcaatcctcctgcctcagccttccaactagctgggactacagg 7021
          |   |||||||| |||||| || |||||||||||||| |||| |||||||| ||||||
Db  46235 CCACCTCCTGGGTTCAAGCGATTCTCCTGCCTCAGCCTCCCAAGTAGCTGGGATTACAGG 46294

Qy   7022 tgcgcgccaccgtgcctggctaacttttcatttttttgtagggacggggtctcgttttg 7081
            | | |||||   |||||||| |||| ||||||| |||| || ||||| ||  | ||
Db  46295 CACCCACCACCACACCTGGCTAATTTTTGTTTTTTTTAGTAGAGATGGGGTTTCACTATG 46354

Qy   7082 ttgccaaagctggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaa 7141
          | | ||| |||||||| ||||| || ||||     |||| ||  | | |||||| ||||
Db  46355 TGGGCAAGGCTGGTCTTGAACTTCTGACCTCGGATGATCCTCCCACTTAGGCCTCTCAAA 46414

Qy   7142 gtgctgagattgcagatgtgagccaccgtgcctggccagatttttctttttattcttcttt 7201
          |||||| |||| |||   ||||| ||||||| |||| ||| ||| ||||| ||    |||
Db  46415 GTGCTGGGATTACAGGCATGAGCAACCGTGCTTGGCGAGAGTTTCCTTTTCTTAGAGTTT 46474

Qy   7202 ctttttctttttgctttcttg 7223
          ||  ||  |  ||  ||   ||
Db  46475 CTGGTTTATCATTTTTTAAATG 46496


Search completed: April  6, 2001, 01:12:41
Job time: 20441 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:21:50 ; Search time 4648.59 Seconds
(without alignments)
13639.351 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: em_estba:*
42: em_estfun:*
43: em_esthum1:*
44: em_esthum2:*
45: em_esthum3:*
46: em_esthum4:*
47: em_esthum5:*
48: em_esthum6:*
49: em_esthum7:*
50: em_esthum8:*
51: em_esthum9:*
52: em_esthum10:*
53: em_esthum11:*
54: em_esthum12:*
55: em_esthum13:*
56: em_esthum14:*
57: em_esthum15:*
58: em_esthum16:*
59: em_esthum17:*
60: em_esthum18:*
61: em_esthum19:*
62: em_esthum20:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estov1:*
68: em_estov2:*
69: em_estpl1:*
70: em_estpl2:*
71: em_estpl3:*
72: em_estpl4:*
73: em_estpl5:*
74: em_estro1:*
75: em_estro2:*
76: em_estro3:*
77: em_estro4:*
78: em_estro5:*
79: em_estro6:*
80: em_estro7:*
81: em_estro8:*
82: em_estro9:*
83: em_estro10:*
84: em_estro11:*
85: em_estro12:*
86: em_estro13:*
87: gb_est41:*
88: gb_est42:*
89: gb_est43:*
90: gb_est44:*
91: gb_est45:*
92: gb_est46:*
93: gb_est47:*
94: gb_est48:*
95: gb_est49:*
96: gb_est50:*
97: gb_est51:*
98: gb_est52:*
99: gb_est53:*
100: gb_est54:*
101: gb_est55:*
102: gb_est56:*
103: gb_est57:*
104: gb_est67:*
105: gb_est68:*
106: gb_est69:*
107: gb_est70:*
108: gb_est71:*
109: gb_est72:*
110: gb_est73:*
111: gb_est74:*
112: em_esthum21:*
113: em_esthum22:*
114: em_esthum23:*
115: em_estom1:*
116: em_estom2:*

117: em_estpl6:*
118: em_estpl7:*
119: em_estpl8:*
120: em_estro14:*
121: em_estro15:*
122: em_estro16:*
123: em_estro17:*
124: em_estro18:*
125: em_estro19:*
126: gb_est58:*
127: gb_est59:*
128: gb_est60:*
129: gb_est61:*
130: gb_est62:*
131: gb_est63:*
132: gb_est64:*
133: gb_est65:*
134: gb_est66:*
135: gb_est75:*
136: gb_est76:*
137: gb_est77:*
138: gb_est78:*
139: gb_est79:*
140: gb_est80:*
141: gb_est81:*
142: gb_est82:*
143: gb_est83:*
144: gb_est84:*
145: gb_est85:*
146: gb_est86:*
147: gb_est87:*
148: gb_gss1:*
149: gb_gss2:*
150: gb_gss3:*
151: gb_gss4:*
152: em_gss1:*
153: em_gss2:*
154: em_gss3:*
155: em_gss4:*
156: gb_gss5:*
157: gb_gss6:*
158: gb_gss7:*
159: gb_gss8:*
160: gb_gss9:*
161: gb_gss10:*
162: gb_gss11:*
163: gb_gss12:*
164: gb_gss13:*
165: gb_gss14:*
166: gb_gss15:*
167: gb_gss16:*
168: gb_gss17:*
169: gb_gss18:*
170: gb_gss19:*
171: gb_gss20:*
172: gb_gss21:*
173: gb_gss22:*
174: gb_gss23:*
175: gb_gss24:*
176: em_gss5:*
177: em_gss6:*
178: em_gss7:*
179: em_gss8:*
180: em_gss9:*
181: em_gss10:*
182: em_gss11:*
183: em_gss12:*
184: em_gss13:*
185: em_gss14:*
186: em_gss15:*
187: em_gss16:*
188: em_gss17:*
189: em_gss18:*
190: gb_gss25:*
191: gb_gss26:*
192: gb_gss27:*
193: gb_gss28:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	704.2	7.8	705	28	AL118597	AL118597 DKFZp761D
c	2	549.8	6.1	599	142	N31948	N31948 yy22g12.s1
c	3	525.2	5.8	547	40	AW137540	AW137540 UI-H-BI1-
c	4	524.4	5.8	545	38	AW007737	AW007737 wt68e06.x
c	5	521.2	5.8	529	92	AW590735	AW590735 hg48h03.x
c	6	502.4	5.6	619	142	N21184	N21184 yx41a10.s1
c	7	484.4	5.4	817	27	AI984777	AI984777 wr85b07.x
c	8	480.8	5.3	743	107	BE410921	BE410921 601303579
c	9	468.8	5.2	472	28	AL135575	AL135575 DKFZp762B
c	10	468.4	5.2	487	19	AI357233	AI357233 qx63a04.x
c	11	468	5.2	673	107	BE384131	BE384131 601272956
c	12	447.2	4.9	452	39	AW051759	AW051759 wx28h07.x
c	13	445.4	4.9	447	15	AI042490	AI042490 ox62e04.x
c	14	440.4	4.9	442	19	AI360882	AI360882 qy01c12.x
c	15	439.8	4.9	453	23	AI633596	AI633596 th70a03.x
c	16	439.4	4.9	451	23	AI636674	AI636674 ts92e04.x
	17	433.6	4.8	643	142	N35845	N35845 yx89f10.r1
c	18	427.8	4.7	432	38	AW002410	AW002410 wu61f05.x
c	19	423.4	4.7	425	134	BE049448	BE049448 xw86f10.x
c	20	421.4	4.7	434	22	AI623326	AI623326 ts80d08.x
c	21	416.4	4.6	418	15	AI078630	AI078630 ox51b09.x
	22	403.6	4.5	495	142	N42784	N42784 yy22g12.r1
	23	401.4	4.4	796	106	BE279166	BE279166 601156630
c	24	392.4	4.3	450	23	AI652496	AI652496 wb29h02.x
c	25	391.8	4.3	401	19	AI362152	AI362152 qy44h07.x
	26	387	4.3	558	107	BE384676	BE384676 601276954
c	27	385	4.3	461	142	N24642	N24642 yx89f10.s1
c	28	368.8	4.1	531	106	BE276168	BE276168 601144424
c	29	365	4.0	397	39	AW028197	AW028197 wv84a01.x
	30	362.4	4.0	797	109	BE540725	BE540725 601065002
c	31	351.4	3.9	463	1	AA020852	AA020852 ze64b08.r
c	32	342.6	3.8	471	141	H09757	H09757 ym01d02.s1
	33	342.6	3.8	751	106	BE312985	BE312985 601150246
c	34	336.6	3.7	344	38	AW016544	AW016544 UI-H-BI0p
c	35	335	3.7	357	15	AI080440	AI080440 ox82c12.s
	36	303	3.3	328	141	H09841	H09841 ym01d02.r1
c	37	300	3.3	323	23	AI636335	AI636335 tz78a05.x
	38	299.4	3.3	302	140	F11822	F11822 HSC32A091 n
c	39	299.2	3.3	349	145	T65388	T65388 yc73b09.s1
	40	299.2	3.3	448	144	R88008	R88008 ym87b02.r1
	41	296.2	3.3	326	28	AL047147	AL047147 DKFZp586E
	42	296	3.3	296	28	AL134288	AL134288 DKFZp547J
c	43	279	3.1	347	144	R87966	R87966 ym87b02.s1
c	44	257.4	2.8	274	140	F09471	F09471 HSC32A092 n
c	45	254.8	2.8	407	38	AV663617	AV663617 AV663617

ALIGNMENTS

RESULT 1
AL118597
LOCUS AL118597 705 bp mRNA EST 29-FEB-2000
DEFINITION DKFZp761D0110_r1 761 (synonym: hamy2) Homo sapiens cDNA clone DKFZp761D0110 5', mRNA sequence.
ACCESSION AL118597
VERSION AL118597.1 GI:5924496
KEYWORDS EST.
SOURCE human.

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 705)
  AUTHORS   Blum,H., Bauersachs,S., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
  TITLE     EST (Blum, et al.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Blum H
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by LMU (Ludwig Maximilians University,
            Munich/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No s1 sequence available.
            This clone (DKFZp761D0110) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES              Location/Qualifiers
     source           1. .705
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="DKFZp761D0110"
                      /clone_lib="761 (synonym: hamy2)"
                      /tissue_type="amygdala"
                      /dev_stage="adult"
                      /lab_host="DH10B"
                      /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      148 a    205 c    203 g    149 t
ORIGIN
```

```
  Query Match             7.8%;  Score 704.2;  DB 28;  Length 705;
  Best Local Similarity   99.7%;  Pred. No. 2.1e-61;
  Matches  703;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   6144 ccagtggcgggtgatggccagcccctggggcccagcccctgtttactggttcttgcaaat 6203
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      1 CCAGTGGCGGGTGATGGCCAGCCCCTGGGGCCCAGCCCCTGTTTACTGGTTCTTGCAAAT 60

Qy   6204 gggagctgagcagcctctggacagccagtgacctttgacctcggtgaccactcttcttta 6263
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 GGGAGCTGAGCAGCCTCTGGACAGCCAGTGACCTTTGACCTCGGTGACCACTCTTCTTTA 120

Qy   6264 agccatagaccctgaggccctgggctgggtgctgggaagggagggttgaaaccaccgtga 6323
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    121 AGCCATAGACCCTGAGGCCCTGGGCTGGGTGCTGGGAAGGGAGGGTTGAAACCACCGTGA 180

Qy   6324 accagagggtgtggctttccagkcaccctcagggagcctccccatctgtccagctggggc 6383
          ||||||||||||||||||||:||||||||||||||||||||||||||||||||||||||
Db    181 ACCAGAGGGTGTGGCTTTCCAGTCACCCTCAGGGAGCCTCCCCATCTGTCCAGCTGGGGC 240

Qy   6384 cagaggctgggagtccctacctgcttcacgttggccggcggctactctggaatgtttttc 6443
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 CAGAGGCTGGGAGTCCCTACCTGCTTCACGTTGGCCGGCGGCTACTCTGGAATGTTTTTC 300

Qy   6444 cctccccagaatcaagcttttgcttgatccagaagagcccatatcactaagatggcatat 6503
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    301 CCTCCCCAGAATCAAGCTTTTGCTTGATCCAGAAGAGCCCATATCACTAAGATGGCATAT 360

Qy   6504 atgtgatctgggcattttcctcctctgcctacagccaggtttagcggcaaaccttttcccc 6563
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    361 ATGTGATCTGGGCATTTTCCTCCTCTGCCTACAGCCAGGTTTAGCGGCAAACCTTTCCCC 420

Qy   6564 cttagcaccttcagggctgagttctgggtttctagaggtcaggacggctcctcagagcgc 6623
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    421 CTTAGCACCTTCAGGGCTGAGTTCTGGGTTTCTAGAGGTCAGGACGGCTCCTCAGAGCGC 480

Qy   6624 caggaagccagagccccaagcaggacgaaaaagaggcatacacacagcagtgtgaatagc 6683
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    481 CAGGAAGCCAGAGCCCCAAGCAGGACGAAAAAGAGGCATACACACAGCAGTGTGAATAGC 540

Qy   6684 ctggccaccagccatcctccctccacctcaagacccccatttgtccsagactaaaggatc 6743
          ||||||||||||||||||||||||||||||||||||||||||||||:||||||||||||||
Db    541 CTGGCCACCAGCCATCCTCCCTCCACCTCAAGACCCCCATTTGTCCGAGACTAAAGGATC 600

Qy   6744 cagagagcagctccctttctcaggagcttgggcagtgccccagggagtccagggtttctc 6803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    601 CAGAGAGCAGCTCCCTTTCTCAGGAGCTTGGGCAGTGCCCCAGGGAGTCCAGGGTTTCTC 660

Qy   6804 tgcagatgtgcggagcgggaggcggtggtagagagagataaaagg 6848
          |||||||||||||||||||||||||||||||||||||||||||||
Db    661 TGCAGATGTGCGGAGCGGGAGGCGGTGGTAGAGAGAGATAAAAGG 705
```

```
RESULT   2
N31948/c
LOCUS       N31948         599 bp    mRNA            EST       10-JAN-1996
DEFINITION  yy22g12.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone
            IMAGE:272038 3', mRNA sequence.
ACCESSION   N31948
VERSION     N31948.1  GI:1152347
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 599)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            High quality sequence stops: 345
            Source: IMAGE Consortium, LLNL
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: m13 -40 forward
            High quality sequence stop: 345.
FEATURES              Location/Qualifiers
     source           1. .599
                      /organism="Homo sapiens"
                      /db_xref="GDB:3881680"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:272038"
                      /clone_lib="Soares melanocyte 2NbHM"
                      /sex="Male"
                      /tissue_type="melanocyte"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Vector: pT7T3D (Pharmacia) with a modified
                      polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                      was primed with a Not I - oligo(dT) primer [5'
                      TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was size selected, ligated to Eco RI
                      adapters (Pharmacia), digested with Not I and cloned into
                      the Not I and Eco RI sites of a modified pT7T3 vector
                      (Pharmacia). Library constructed by Bento Soares and
                      M.Fatima Bonaldo. RNA from normal foreskin melanocytes
                      (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT      159 a    162 c    159 g    116 t      3 others
ORIGIN
```

```
  Query Match             6.1%;  Score 549.8;  DB 142;  Length 599;
  Best Local Similarity   97.8%;  Pred. No. 4e-46;
  Matches  587;  Conservative    0;  Mismatches   10;  Indels    3;  Gaps    3;
```

```
Qy   8446 tcaccaaggatcggactc-tgcccctggacctgggaacgactggactgtcacggggttcc 8504
          ||||||||||||||||  ||||| |||||||||||||||||||||||||| |||||||
Db    599 TCACCAAGGATCGGACTNTTGCCCTTGGACCTGGGAACGACTGGACTGTCAC-GGGTTCC 541

Qy   8505 ctcctagctctcccagtgaactcctgccaggcacacacagcccctatagcactgagctca 8564
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    540 NTCCTAGCTCTCCCAGTGAACTCCTGCCAGGCACACACAGCCCCTATAGCACTGAGCTCA 481

Qy   8565 catgggactgggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtg 8624
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    480 CATGGGACTGGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTG 421

Qy   8625 cctggccccagtctgggccatctcttaggtgagacagttgcccgaaact-aagccaggcc 8683
          ||  ||||||| ||||||||||||||||||||||||||||||||||| ||||||||||
Db    420 CCCTGCCCCAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAAGCCAGGCC 361

Qy   8684 tggctggaggagcagcagcttggggagagggatttccctgcagacctcaagccatcatgc 8743
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    360 TGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGC 301

Qy   8744 ggtgggtgctgccatgacagaggctgcacccctgggccagcggggctgctcacccacctc 8803
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    300 GGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTC 241

Qy   8804 ttgtgcaaggtggccctttgtgctgcgcctgcaggcagagctggagcccccagcagaggca 8863
          ||||||||||||||||    ||||||||||||||||||||||||||||||||||||||||
Db    240 TTGTGCAAGGTGGCCTTTTGTNTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCA 181

Qy   8864 ggctgggacggaccagcatctggaagatgtacatagttatttttctctttgtggtttctt 8923
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    180 GGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTT 121

Qy   8924 gtttggtttggtttgcttttgacagcttcattttatttttgacgtcacttttttggccatg 8983
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    120 GTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATG 61

Qy   8984 taaactatttgtggcaattttatgttttttatttatgaataaagaatgccatttctcacgc 9043
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     60 TAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACGC 1
```

RESULT 3
AW137540/c
LOCUS AW137540 547 bp mRNA EST 29-OCT-1999
DEFINITION UI-H-BI1-acv-h-12-0-UI.s1 NCI_CGAP_Sub3 Homo sapiens cDNA clone IMAGE:2715887 3', mRNA sequence.
ACCESSION AW137540
VERSION AW137540.1 GI:6141858
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 547)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. cDNA Library Preparation: M.B. Soares Lab Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M13 Forward
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .547
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2715887"
/clone_lib="NCI_CGAP_Sub3"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; The NCI_CGAP_Sub3 library is a subtracted library derived from the NCI_CGAP_Sub1 library, which is a subtracted library derived from BI. BI constitutes a mixture of 21 normalized or subtracted NCI_CGAP libraries: NCI_CGAP_Co4, NCI_CGAP_Pr22, NCI_CGAP_Pr28, NCI_CGAP_Co10, NCI_CGAP_Co16, NCI_CGAP_Kid5, NCI_CGAP_Kid12, NCI_CGAP_Kid3, NCI_CGAP_Kid11, NCI_CGAP_Lym2, NCI_CGAP_Br2, NCI_CGAP_Co8, NCI_CGAP_CLL1, NCI_CGAP_Lei2, NCI_CGAP_Brn23, NCI_CGAP_Lu5, NCI_CGAP_Lu24, NCI_CGAP_Lu19, NCI_CGAP_GC4, NCI_CGAP_GC6, NCI_CGAP_Brn25. These 21 libraries were pooled and a single-stranded DNA preparation of the resulting mixture was used as a tracer in a subtractive hybridization with a driver whose composition is detailed below: NCI_CGAP_Kid3 pool 1 LLAM 3334-3337, 3682-3683, 3798-3803 (IMAGE CloneIDs 1322376-1323911, 1456008-1456775, 1500552-1502855); NCI_CGAP_Kid5 pool 1 LLAM 3338-3342, 3722-3725, 3776-3778 (IMAGE CloneIDs 1323912-1325831, 1471368-1472903, 1492104-1493255); NCI_CGAP_Lu5 pool 1 LLAM 3575-3582, 3851-3854 (IMAGE CloneIDs 1414920-1417991, 1520904-1522439); NCI_CGAP_GC4 pool 1 LLAM 3164-3167, 3716-3720, 3733-3735 (IMAGE CloneIDs 1257096-1258631, 1469064-1470983, 1475592-1476743); NCI_CGAP_Pr22 pool 1 LLAM 2457-2459, 2758-2759, 3062-3068 (IMAGE CloneIDs 985608-986759, 1101192-1101959, 1217928-1220615); NCI_CGAP_Co10 pool 1 LLAM 2644-2653, 2871-2872 (IMAGE CloneIDs 1057416-1061255, 1144584-1145351). Subtraction was performed as previously described [Bonaldo, Lennon & Soares (1996): Normalization and Subtraction: Two Approaches To Facilitate Gene Discovery. Genome Research 6, 791-806.
TAG_LIB=NCI_CGAP_Brn23
TAG_TISSUE=brain
TAG_SEQ=ATATC"
BASE COUNT 143 a 147 c 138 g 119 t
ORIGIN

Query Match 5.8%; Score 525.2; DB 40; Length 547;
Best Local Similarity 99.4%; Pred. No. 1.1e-43;
Matches 527; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   8513 tctcccagtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggac 8572
          |||||| |||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db     547 TCTCCCCGTGAACTCGTGCCAGGCACACACAGCCCCTATAGCACTGAGCTCACATGGGAC 488

Qy   8573 tgggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggccc 8632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     487 TGGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCC 428

Qy   8633 cagtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggag 8692
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     427 CAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAG 368

Qy   8693 gagcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgc 8752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     367 GAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGC 308

Qy   8753 tgccatgacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaag 8812
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     307 TGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGCAAG 248

Qy   8813 gtggcctttgtgctgcgcctgcaggcagagctggagcccccagcagaggcaggctgggac 8872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     247 GTGGCCTTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGAC 188

Qy   8873 ggaccagcatctggaagatgtacatagttatttttctctttgtggtttcttgtttggttt 8932
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 GGACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTGTTTGGTTT 128

Qy   8933 ggtttgctttttgacagcttcattttatttttgacgtcacttttttggccatgtaaactatt 8992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 GGTTTGCTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATT 68

Qy   8993 tgtggcaatttttatgtttttatttatgaataaagaatgccatttctcacg 9042
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 TGTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACG 18


RESULT   4
AW007737/c
LOCUS       AW007737      545 bp    mRNA            EST       09-MAR-2000
DEFINITION  wt68e06.x1 Soares_thymus_NHFTh Homo sapiens cDNA clone
            IMAGE:2512642 3' similar to contains MER22.t3 TAR1 repetitive
            element ;, mRNA sequence.
ACCESSION   AW007737
VERSION     AW007737.1  GI:5856515
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 545)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 1333   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 463.
FEATURES             Location/Qualifiers
     source          1. .545
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2512642"
                     /clone_lib="Soares_thymus_NHFTh"
                     /dev_stage="fetal"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: thymus, pooled; Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
                     oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCAACGTTTTTTTTTTTTTTTTT 3'],
                     double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector.  Library
                     went through one round of normalization. Library
                     constructed by Bento Soares and M. Fatima Bonaldo.  "
BASE COUNT       80 a    159 c    189 g    117 t
ORIGIN


  Query Match             5.8%;  Score 524.4;  DB 38;  Length 545;
  Best Local Similarity   99.8%;  Pred. No. 1.4e-43;
  Matches  525;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    526 GGGCTCCGAGAAGGGTGCAGTGAGGCCCACAGCCTTCAAGCCTGTGCTGCCACGGTCAGG 467

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    466 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGA 407

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    406 CAAGCCCAAGGAGCAGGAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 347

Qy   1826 ccggaactccatgtccagcctgcccacacacagcaccagcagcagctaccagctggaccc 1885
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    346 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCC 287

Qy   1886 gctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcaccca 1945
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    286 GCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCA 227

Qy   1946 gggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccgacgg 2005
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    226 GGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGG 167

Qy   2006 aggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctcccc 2065
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    166 AGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCC 107

Qy   2066 catctccacggacgagtgcagcatccaggagctggaacagaagctgttggagagggaggg 2125
          |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    106 CATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGG 47

Qy   2126 cgccctccagaagctgcagcgcagctttgaggagaaggagcttgcc 2171
          |||||||||||||||||||||||||||||||||||||||||||||
Db     46 CGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCC 1


RESULT   5
AW590735/c
LOCUS       AW590735      529 bp    mRNA            EST       22-MAR-2000
DEFINITION  hg48h03.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:2948885 3',
            mRNA sequence.
ACCESSION   AW590735
VERSION     AW590735.1  GI:7277875
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 529)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            R. Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            image.llnl.gov/image/html/iresources.shtml
            Seq primer: -40UP from Gibco
            High quality sequence stop: 469.
FEATURES             Location/Qualifiers
     source          1. .529
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2948885"
                     /clone_lib="NCI_CGAP_GC6"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Plasmid DNA from the normalized library
                     NCI_CGAP_GC4 was prepared, and ss circles were made in
                     vitro. Following HAP purification, this DNA was used as
                     tracer in a subtractive hybridization reaction. The driver
                     was PCR-amplified cDNAs from a pool of 5,000 clones made
```

from the same library (cloneIDs 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo. "
BASE COUNT 141 a 146 c 134 g 108 t
ORIGIN

Query Match 5.8%; Score 521.2; DB 92; Length 529;
Best Local Similarity 99.4%; Pred. No. 2.9e-43;
Matches 523; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```
Qy   8519 agtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggactgggat 8578
          ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    529 AGTGAACTCATGCCAGGCACACACAGCCCCTATAGCACTGAGCTCACATGGGACTGGGAT 470

Qy   8579 atgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggccccagtct 8638
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| ||
Db    469 ATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCCCAGCCT 410

Qy   8639 gggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggaggagcag 8698
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    409 GGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAGGAGCAG 350

Qy   8699 cagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgctgccat 8758
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    349 CAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGCTGCCAT 290

Qy   8759 gacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaaggtggcc 8818
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 GACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGCAAGGTGGCC 230

Qy   8819 tttgtgctgcgcctgcaggcagagctggagcccccagcagaggcaggctgggacggacca 8878
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 TTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGACGGACCA 170

Qy   8879 gcatctggaagatgtacatagttatttttctctttgtggtttcttgtttggtttggtttg 8938
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 GCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTG 110

Qy   8939 cttttgacagcttcattttatttttgacgtcacttttttggccatgtaaactatttgtggc 8998
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    109 CTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATTTGTGGC 50

Qy   8999 aattttatgtttttatttatgaataaagaatgccatttctcacgcc 9044
          |||||||||||||||||||||||||||||||||||||||||||| |
Db     49 AATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACGAC 4
```

RESULT 6
N21184/c
LOCUS N21184 619 bp mRNA EST 19-DEC-1995
DEFINITION yx41a10.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:264282 3', mRNA sequence.
ACCESSION N21184
VERSION N21184.1 GI:1126354
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 619)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 342
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2377 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 342.
FEATURES Location/Qualifiers
source 1. .619
/organism="Homo sapiens"
/db_xref="GDB:3873924"
/db_xref="taxon:9606"
/clone="IMAGE:264282"
/clone_lib="Soares melanocyte 2NbHM"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT 165 a 163 c 160 g 125 t 6 others
ORIGIN

Query Match 5.6%; Score 502.4; DB 142; Length 619;
Best Local Similarity 93.5%; Pred. No. 1.9e-41;
Matches 576; Conservative 0; Mismatches 32; Indels 8; Gaps 5;

```
Qy   8427 agcagcactccgcatcacctcaccaaggatcggactctgcccctggacctgggaacgact 8486
          ||||||| || | || || |||||| |    |   |    |||  ||| ||||||||||||
Db    608 AGCAGCANTCGGAATAACTTCACCAGGATCGGATNT---GCCCCTGACNTGGGAACGACT 552

Qy   8487 ggactgtcacggggttccctcctagctctcccagtgaactcctgccaggcacacacagcc 8546
          ||||| |||| |||||||||||||||| | ||  |||||||||||||||||||||||||||
Db    551 GGACTTTCAC-GGGTTCCCTCCTAGCTTTTCCNATGAACTCCTGCCAGGCACACACAGCC 493

Qy   8547 cctatagcactgagctcacatgggactgggatatgggggcatctcttccccagagaggca 8606
          |||||||||||||  ||||||||||||||      ||||||||||||||||||||||||||
Db    492 CCTATAGCACTGACTTCACATGGGACTGG--ATATGGGGCATCTCTTCCCCAGAGAGGCA 435

Qy   8607 ctcagtgagcctcctgtgcctggccccagtctgggccatctcttaggtgagacagttgcc 8666
          |||||||||| |||||||||| ||||||| ||||||||||||||||||||||||||||||
Db    434 CTCAGTGAGCNTCCTGTGCCT-GCCCCAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCC 376

Qy   8667 cgaaactaagccaggcctggctggaggagcagcagcttggggagagggatttccctgcag 8726
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 CGAAACTAAGCAAGGCCTGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAG 316

Qy   8727 acctcaagccatcatgcggtgggtgctgccatgacagaggctgcacccctgggccagcgg 8786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 ACCTCAAGCCATCATGCGGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCCAGCGG 256

Qy   8787 ggctgctcacccacctcttgtgcaaggtggcctttgtgctgcgcctgcaggcagagctgg 8846
          |||||||||||||||||||||||||||||||||||  ||||||||||||||||||||||
Db    255 GGCTGCTCACCCACCTCTTGTGCAAGGTGGCCTTTGT-NTGCGCCTGCAGGCAGAGCTGG 197

Qy   8847 agccccccagcagaggcaggctgggacggaccagcatctggaagatgtacatagttatttt 8906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 AGCCCCCAGCAGAGGCAGGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTTATTTT 137

Qy   8907 tctctttgtggtttcttgtttggtttggtttgcttttgacagcttcattttatttttgac 8966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 TCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTTTTGAC 77

Qy   8967 gtcacttttttggccatgtaaactatttgtggcaattttatgttttttatttatgaataaag 9026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db     76 GTCACTTTTTGGCCATGTAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAATAAAG 17

Qy   9027 aatgccatttctcacg 9042
          ||||||||||||||||
Db     16 AATGCCATTTCTCACG 1


RESULT   7
AI984777/c
LOCUS       AI984777      817 bp    mRNA            EST       08-MAR-2000
DEFINITION  wr85b07.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2494453 3'
            similar to TR:O60299 O60299 KIAA0552 PROTEIN. ;, mRNA sequence.
ACCESSION   AI984777
VERSION     AI984777.1  GI:5812054
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 817)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 1307   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 428.
FEATURES             Location/Qualifiers
     source          1. .817
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2494453"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo. "
BASE COUNT      133 a    235 c    266 g    180 t      3 others
ORIGIN


  Query Match             5.4%;  Score 484.4;  DB 27;  Length 817;
  Best Local Similarity   97.2%;  Pred. No. 1e-39;
  Matches  514;  Conservative    0;  Mismatches   12;  Indels    3;  Gaps    2;

Qy   1646 gggctccgagaagggtgcagtgaggcccacagccttcaagcctgtgctgccacggtcagg 1705
          ||||||  |||||||||||||| |||||||||||||||||| ||||||||||||||||||
Db    529 GGGCTCGTAGAAGGGTGCAGTAAGGCCCACAGCCTTCAAGCTTGTGCTGCCACGGTCAGG 470

Qy   1706 agccatcctgcactcctccccggagagtgccagccaccagctgcaccccgcccctccaga 1765
          |||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||
Db    469 AGCCATCCTGCACTCCTCCCCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCNTCCAGA 410

Qy   1766 caagcccaaggagcaggagctgaagcctggcctgtgctctggggcgctgtcagactccgg 1825
          |||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||

Db    409 CAAGCCCAAGGAGCAGAAGCTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGG 350

Qy   1826 ccggaactccatgtccagcctgcccacacacag-caccagcagcagctaccagct--gga 1882
          |||||||||||||||||||||||||||||||| | ||||||| || |||||  |  |||
Db    349 CCGGAACTCCATGTCCAGCCTGCCCACACACAGCCCCCAGCAGAAGTTACCAAGTTGGGA 290

Qy   1883 cccgctggtcacacccgtgggacccacaagccgttttgggggctccgcccacaacatcac 1942
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    289 CCCGCTGGTCACACCCGTGGGACCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCAC 230

Qy   1943 ccagggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccga 2002
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    229 CCAGGGCATCGTCCTCCAGGACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGA 170

Qy   2003 cggaggtagcaagctgggccactcgaacaaggcagacaagggcccctcgtgtgtccgctc 2062
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    169 CGGAGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTC 110

Qy   2063 ccccatctccacggacgagtgcagcatccaggagctggaacagaagctgttggagaggga 2122
          |||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    109 CCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGA 50

Qy   2123 gggcgccctccagaagctgcagcgcagctttgaggagaaggagcttgcc 2171
          |||||||||||||||||||||||||||||||||||||||||||||||||
Db     49 GGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCC 1


RESULT   8
BE410921/c
LOCUS       BE410921      743 bp    mRNA            EST       21-JUL-2000
DEFINITION  601303579F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:3638203 5',
            mRNA sequence.
ACCESSION   BE410921
VERSION     BE410921.1  GI:9347371
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 743)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM338  row: m  column: 20
            High quality sequence start: 49
            High quality sequence stop: 688.
FEATURES             Location/Qualifiers
     source          1. .743
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3638203"
                     /clone_lib="NIH_MGC_21"
                     /tissue_type="choriocarcinoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: placenta; Vector: pOTB7; Site_1: XhoI;
                     Site_2: EcoRI; cDNA made by oligo-dT priming.
                     Directionally cloned into EcoRI/XhoI sites using the
                     following 5' adaptor: GGCACGAG(G).  Size-selected >500bp
                     for average insert size 1.8kb.  Library constructed by
                     Ling Hong in the laboratory of Gerald M. Rubin (University
                     of California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      119 a    229 c    226 g    168 t      1 others
ORIGIN
```

```
  Query Match             5.3%;  Score 480.8;  DB 107;  Length 743;
  Best Local Similarity   81.1%;  Pred. No. 2.4e-39;
  Matches  658;  Conservative    0;  Mismatches   13;  Indels  140;  Gaps    3;

Qy   1923 ggctccgcccacaacatcacccagggcatcgtcctccaggacagcaacatgatgagcctg 1982
          ||  |||||||| | ||||| |||| |||||||||||||||| ||||||||||||||||
Db    711 GGGCCCGCCCACGAAATCACACAGG-CATCGTCCTCCAGGACNGCAACATGATGAGCCTG 653

Qy   1983 aaggctctgtccttctccgacggaggtagcaagctgggccactcgaacaaggcagacaag 2042
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    652 AA-GCTCTGTCCTTCTCCGACGGAGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAG 594

Qy   2043 ggcccctcgtgtgtccgctcccccatctccacggacgagtgcagcatccaggagctggaa 2102
          |||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db    593 GGCCCCTCGTGTGTCCGCTCCCCCATCGCCACGGACGAGTGCAGCATCCAGGAGCTGGAG 534

Qy   2103 cagaagctgttggagagggagggcgccctccagaagctgcagcgcagctttgaggagaag 2162
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    533 CAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAG 474

Qy   2163 gagcttgcctccagcctggcctacgaggagcggccgcggcgctgcagggacgagctggag 2222
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    473 GAGCTTGCCTCCAGCCTGGCCTACGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAG 414

Qy   2223 ggcccggagcccaaaggcggcaacaagctcaagcaggcctcgcagaagagccagcgcgcg 2282
          ||||||||||||||||||||||||||||||
Db    413 GGCCCGGAGCCCAAAGGCGGCAACAAGCTCAA---------------------------- 380

Qy   2283 cagcaggtcctgcacctgcaggtactgcagcttcagcaggagaagcggcagctccggcag 2342

Db    381 ------------------------------------------------------------ 380

Qy   2343 gagctcgagagcctcatgaaggagcaggacctgctggagaccaagctcaggtcctacgag 2402
                                                          ||||||||||
Db    381 ----------------------------------------------------GTCCTACGAG 372

Qy   2403 agggagaagaccagcttcggccccgcgctggaggagacccagtgggaggtgaggccacac 2462
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    371 AGGGAGAAGACCAGCTTCGGCCCCGCGCTGGAGGAGACCCAGTGGGAGGTGAGGCCACAC 312

Qy   2463 agggctcatgggtttgggtggtcagcggtttggcgccagtacccccctctccttctggtg 2522
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    311 AGGGCTCATGGGTTTGGGTGGTCAGCGGTTTGGCGCCAGTACCCCCCTCTCCTTCTGGTG 252

Qy   2523 ctggccaatagcgtgcaaacacagaccgcgcaggcaagcggggctaatgtgctggcttta 2582
          | ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    251 CCGGCCAATAGCGTGCAAACACAGACCGCGCAGGCAAGCGGGGCTAATGTGCTGGCTTTA 192

Qy   2583 tcacccaaagaaggggctccctgcaaaccatgttgggggatcgacttacatctgagcttc 2642
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    191 TCACCCAAAGAAGGGGCTCCCTGCAAACCATGTTGGGGGATCGACTTACATCTGAGCTTC 132

Qy   2643 ctcctgtccccaccatcaccctcatggctcctagatttcagtttcccaagtgagccatta 2702
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    131 CTCCTGTCCCCACCATCACCCTCATGGCTCCTAGATTTCAGTTTCCCAAGTGAGCCATTA 72

Qy   2703 aatcatgaagccggaagccagatgaccaagg 2733
          |||||||||||||||||||||||  | | |
Db     71 AATCATGAAGCCGGAAGCCAGATGTGCCAAG 41


RESULT   9
AL135575/c
LOCUS       AL135575      472 bp    mRNA            EST       25-FEB-2000
DEFINITION  DKFZp762B204_r1 762 (synonym: hmel2) Homo sapiens cDNA clone
            DKFZp762B204 5', mRNA sequence.
ACCESSION   AL135575
VERSION     AL135575.1  GI:6603762
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 472)
  AUTHORS   Koehrer,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.
  TITLE     EST (Koehrer,K., Beyer,A., Mewes,H.W., Weil,B. and Wiemann,S.)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Koehrer K
            MIPS
            Am Klopferspitz 18a D-82152 Martinsried, Germany
            This is the 5' sequence of the clone insert
            Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
            Research Center (DKFZ); Email s.wiemann@dkfz- heidelberg.de;
            sequenced by BMFZ (Biomedical Research Center at the Charite,
            Berlin/Germany) within the cDNA sequencing consortium of the German
            Genome Project.
            No sl sequence available.
            This clone (DKFZp762B204) is available at the RZPD in Berlin.
            Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
            Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES             Location/Qualifiers
     source          1. .472
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="DKFZp762B204"
                     /clone_lib="762 (synonym: hmel2)"
                     /tissue_type="melanoma (MeWo cell line)"
                     /dev_stage="adult"
                     /lab_host="DH10B"
                     /note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
BASE COUNT      133 a    134 c    117 g     88 t
ORIGIN


  Query Match             5.2%;  Score 468.8;  DB 28;  Length 472;
  Best Local Similarity   99.6%;  Pred. No. 4.6e-38;
  Matches  470;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy   8574 gggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggcccc 8633
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    472 GGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCCC 413

Qy   8634 agtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggagg 8693
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    412 AGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAGG 353

Qy   8694 agcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgct 8753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    352 AGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGCT 293

Qy   8754 gccatgacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaagg 8813
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    292 GCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGCAAGG 233

Qy   8814 tggcctttgtgctgcgcctgcaggcagagctggagcccccagcagaggcaggctgggacg 8873
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    232 TGGCCTTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGACG 173

Qy   8874 gaccagcatctggaagatgtacatagttatttttctctttgtggtttcttgtttggtttg 8933
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    172 GACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTGTTTGGTTTG 113

Qy   8934 gtttgcttttgacagcttcattttatttttgacgtcacttttttggccatgtaaactattt 8993
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db    112 GTTTGCTTTTGACAGTTTCATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATTT 53

Qy   8994 gtggcaatttttatgtttttatttatgaataaagaatgccatttctcacgccc 9045
          ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     52 GTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACGCCC 1


RESULT  10
AI357233/c
LOCUS       AI357233      487 bp    mRNA            EST       15-FEB-1999
```

```
DEFINITION  qx63a04.x1 NCI_CGAP_GC4 Homo sapiens cDNA clone IMAGE:2005998 3',
            mRNA sequence.
ACCESSION   AI357233
VERSION     AI357233.1  GI:4108854
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 487)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 743   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 444.
FEATURES             Location/Qualifiers
     source          1. .487
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2005998"
                     /clone_lib="NCI_CGAP_GC4"
                     /tissue_type="pooled germ cell tumors"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; 1st strand cDNA was prepared from 3 pooled
                     germ cell tumors, and was then primed with a Not I -
                     oligo(dT) primer. Double-stranded cDNA was ligated to Eco
                     RI adaptors (Pharmacia), digested with Not I and cloned
                     into the Not I and Eco RI sites of the modified pT7T3
                     vector.  Library is normalized. Library was constructed by
                     Bento Soares and M. Fatima Bonaldo. "
BASE COUNT      133 a    134 c    116 g    104 t
ORIGIN
```

```
  Query Match             5.2%;  Score 468.4;  DB 19;  Length 487;
  Best Local Similarity   99.8%;  Pred. No. 4.9e-38;
  Matches  469;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   8573 tgggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggccc 8632
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    487 TGGGATATGGGGGCATCTCTTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCC 428

Qy   8633 cagtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggag 8692
          ||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    427 CAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAG 368

Qy   8693 gagcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgc 8752
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    367 GAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGC 308

Qy   8753 tgccatgacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaag 8812
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    307 TGCCATGACAGAGGCTGCACCCCTGGGCCAGCGGGGCTGCTCACCCACCTCTTGTGCAAG 248

Qy   8813 gtggcctttgtgctgcgcctgcaggcagagctggagcccccagcagaggcaggctgggac 8872
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    247 GTGGCCTTTGTGCTGCGCCTGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGAC 188

Qy   8873 ggaccagcatctggaagatgtacatagttatttttctctttgtggtttcttgtttggttt 8932
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    187 GGACCAGCATCTGGAAGATGTACATAGTTATTTTTCTCTTTGTGGTTTCTTGTTTGGTTT 128

Qy   8933 ggtttgcttttgacagcttcatttttatttttgacgtcacttttttggccatgtaaactatt 8992
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    127 GGTTTGCTTTTGACAGCTTCATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATT 68

Qy   8993 tgtggcaattttatgtttttatttatgaataaagaatgccatttctcacg 9042
          ||||||||||||||||||||||||||||||||||||||||||||||||||
Db     67 TGTGGCAATTTTATGTTTTTATTTATGAATAAAGAATGCCATTTCTCACG 18
```

```
RESULT  11
BE384131/c
LOCUS       BE384131      673 bp    mRNA            EST       21-JUL-2000
DEFINITION  601272956F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:3614076 5',
            mRNA sequence.
ACCESSION   BE384131
VERSION     BE384131.1  GI:9329496
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 673)
  AUTHORS   NIH-MGC http://www.ncbi.nlm.nih.gov/MGC/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: ATCC/DCTD/DTP
             cDNA Library Preparation: Ling Hong/Rubin Laboratory
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Incyte Genomics, Inc.
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
            Plate: LLCM275  row: p  column: 13
            High quality sequence stop: 670.
FEATURES             Location/Qualifiers
     source          1. .673
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:3614076"
                     /clone_lib="NIH_MGC_20"
                     /tissue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT      110 a    211 c    208 g    144 t
ORIGIN
```

```
  Query Match             5.2%;  Score 468;  DB 107;  Length 673;
  Best Local Similarity   81.9%;  Pred. No. 4.6e-38;
  Matches  663;  Conservative    0;  Mismatches    5;  Indels  142;  Gaps    5;

Qy   1943 ccagggcatcgtcctccaggacagcaacatgatgagcctgaaggctctgtccttctccga 2002
          ||  ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    669 CCCAGGCATCGTCCTCCA-GACAGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGA 611

Qy   2003 cggaggtagcaagctgggccactcgaacaaggcagacaaggg-cccctcgtgtgtccgct 2061
          |||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    610 CGGAGGTAGCAAGCTGGGCCACTCGAACAAGGCAGACAAGGGCCCCCTCGTGTGTCCGCT 551

Qy   2062 cccccatctccacggacgagtgcagcatccaggagctggaacagaagctgttggagaggg 2121
          |||||||||||||||||||||||||||||||||||||||| ||||||||||||||||||
```

```
Db    550 CCCCCATCTCCACGGACGAGTGCAGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGG 491

Qy   2122 agggcgccctccagaagctgcagcgcagctttgaggagaaggagcttgcctccagcctgg 2181
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    490 AGGGCGCCCTCCAGAAGCTGCAGCGCAGCTTTGAGGAGAAGGAGCTTGCCTCCAGCCTGG 431

Qy   2182 cctacgaggagcggccgcggcgctgcagggacgagctggagggcccggagcccaaaggcg 2241
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    430 CCTACGAGGAGCGGCCGCGGCGCTGCAGGGACGAGCTGGAGGGCCCGGAGCCCAAAGGCG 371

Qy   2242 gcaacaagctcaagcaggcctcgcagaagagccagcgcgcgcagcaggtcctgcacctgc 2301
          ||||||||||||
Db    370 GCAACAAGCTCAA---------------------------------------------- 356

Qy   2302 aggtactgcagcttcagcaggagaagcggcagctccggcaggagctcgagagcctcatga 2361

Db    357 ---------------------------------------------------------- 356

Qy   2362 aggagcaggacctgctggagaccaagctcaggtcctacgagagggagaagaccagcttcg 2421
                                        ||||||||||||||||||||||||||||
Db    357 -------------------------------GTCCTACGAGAGGGAGAAGACCAGCTTCG 329

Qy   2422 gccccgcgctggaggagacccagtgggaggtgaggccacacagggctcatgggtttgggt 2481
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    328 GCCCCGCGCTGGAGGAGACCCAGTGGGAGGTGAGGCCACACAGGGCTCATGGGTTTGGGT 269

Qy   2482 ggtcagcggtttggcgccagtacccccctctccttctggtgctggccaatagcgtgcaaa 2541
          ||||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||
Db    268 GGTCAGCGGTTTGGCGCCAGTACCCCCCTCTCCTTCTGGTGCCGGCCAATAGCGTGCAAA 209

Qy   2542 cacagaccgcgcaggcaagcggggctaatgtgctggctttatcacccaaagaaggggctc 2601
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    208 CACAGACCGCGCAGGCAAGCGGGGCTAATGTGCTGGCTTTATCACCCAAAGAAGGGGCTC 149

Qy   2602 cctgcaaaccatgttgggggatcgacttacatctgagcttcctcctgtccccaccatcac 2661
          ||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    148 CCTGC-AACCATGTTGGGGGATCGACTTACATCTGAGCTTCCTCCTGTCCCCACCATCAC 90

Qy   2662 cctcatggctcctagatttcagtttcccaagtgagccattaaatcatgaagccggaagcc 2721
          |||||||||||||||||||||||||| ||||||||||||||||||||||||||||||||
Db     89 CCTCATGGCTCCTAGATTTCAGTTTCCCACGTGAGCCATTAAATCATGAAGCCGGAAGCC 30

Qy   2722 agatgaccaaggcccagccaggctgtgggc 2751
          |||||||| ||||||||||||||||||||
Db     29 AGATGACC-AGGCCCAGCCAGGCTGTGGGC 1


RESULT 12
AW051759/c
LOCUS       AW051759      452 bp    mRNA            EST       09-MAR-2000
DEFINITION  wx28h07.x1 NCI_CGAP_Kid11 Homo sapiens cDNA clone IMAGE:2545021 3',
            mRNA sequence.
ACCESSION   AW051759
VERSION     AW051759.1  GI:5914118
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 452)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 610    Std Error: 0.00
            Seq primer: -40UP from Gibco.
FEATURES             Location/Qualifiers
     source          1. .452
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2545021"
                     /clone_lib="NCI_CGAP_Kid11"
                     /lab_host="DH10B"
                     /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
                     a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                     Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
                     prepared, and ss circles were made in vitro. Following HAP
                     purification, this DNA was used as tracer in a subtractive
                     hybridization reaction. The driver was PCR-amplified cDNAs
                     from a pool of 5,000 clones made from the same library
                     (cloneIDs 1322376-1323911, 1456007-1456775, and
                     1500552-1502855). Subtraction by Bento Soares and M.
                     Fatima Bonaldo.  "
BASE COUNT      128 a    126 c    112 g     86 t
ORIGIN


  Query Match             4.9%;  Score 447.2;  DB 39;  Length 452;
  Best Local Similarity   99.3%;  Pred. No. 6.3e-36;
  Matches  449;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy   8592 ttccccagagaggcactcagtgagcctcctgtgcctggccccagtctgggccatctctta 8651
          |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||||||
Db    452 TTCCCCAGAGAGGCACTCAGTGAGCCTCCTGTGCCTGGCCCCAGCCTGGGCCATCTCTTA 393

Qy   8652 ggtgagacagttgcccgaaactaagccaggcctggctggaggagcagcagcttggggaga 8711
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    392 GGTGAGACAGTTGCCCGAAACTAAGCCAGGCCTGGCTGGAGGAGCAGCAGCTTGGGGAGA 333

Qy   8712 gggatttccctgcagacctcaagccatcatgcggtgggtgctgccatgacagaggctgca 8771
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    332 GGGATTTCCCTGCAGACCTCAAGCCATCATGCGGTGGGTGCTGCCATGACAGAGGCTGCA 273

Qy   8772 cccctgggccagcggggctgctcacccacctcttgtgcaaggtggcctttgtgctgcgcc 8831
          |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Db    272 CCCCTGGGCCAGCGGGGCTGCTCACCCACTTCTTGTGCAAGGTGGCCTTTGTGCTGCGCC 213

Qy   8832 tgcaggcagagctggagcccccagcagaggcaggctgggacggaccagcatctggaagat 8891
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    212 TGCAGGCAGAGCTGGAGCCCCCAGCAGAGGCAGGCTGGGACGGACCAGCATCTGGAAGAT 153

Qy   8892 gtacatagttatttttctctttgtggtttcttgtttggtttggtttgcttttgacagctt 8951
          ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    152 GTACATAGTTATTTATCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCTTTTGACAGCTT 93

Qy   8952 cattttatttttgacgtcacttttttggccatgtaaactatttgtggcaatttttatgtttt 9011
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     92 CATTTTATTTTTGACGTCACTTTTTGGCCATGTAAACTATTTGTGGCAATTTTATGTTTT 33

Qy   9012 tatttatgaataaagaatgccatttctcacgc 9043
          |||||||||||||||||||||||||||||||
Db     32 TATTTATGAATAAAGAATGCCATTTCTCACGC 1


RESULT 13
AI042490/c
LOCUS       AI042490      447 bp    mRNA            EST       28-AUG-1998
DEFINITION  ox62e04.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:1660926
            3', mRNA sequence.
ACCESSION   AI042490
VERSION     AI042490.1  GI:3281684
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 447)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Insert Length: 656   Std Error: 0.00
            Seq primer: -40m13 fwd. ET from Amersham
            High quality sequence stop: 435.
FEATURES             Location/Qualifiers
     source          1. .447
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:1660926"
                     /clone_lib="Soares_NhHMPu_S1"
                     /tissue_type="Pooled human melanocyte, fetal heart, and
                     pregnant uterus"
                     /lab_host="DH10B"
                     /note="Organ: mixed (see below); Vector: pT7T3D-Pac
                     (Pharmacia) with a modified polylinker; Site_1: Not I;
                     Site_2: Eco RI; Equal amounts of plasmid DNA from three
                     normalized libraries (melanocyte 2NbHM, pregnant uterus
                     NbHPU, and fetal heart NbHH19W) were mixed, and ss circles
                     were made in vitro. Following HAP purification, this DNA
                     was used as tracer in a subtractive hybridization
                     reaction. The driver was PCR-amplified cDNAs from pools of
                     5,000 clones made from the same 3 libraries. The pools
                     consisted of I.M.A.G.E. clones 260232-265223,
                     340488-345479, and 484488-489479."
BASE COUNT       62 a    133 c    156 g     96 t
ORIGIN


  Query Match             4.9%;  Score 445.4;  DB 15;  Length 447;
  Best Local Similarity   99.8%;  Pred. No. 9.6e-36;
  Matches  446;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    1725 ccggagagtgccagccaccagctgcaccccgcccctccagacaagcccaaggagcaggag 1784
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     447 CCGGAGAGTGCCAGCCACCAGCTGCACCCCGCCCCTCCAGACAAGCCCAAGGAGCAGGAG 388

Qy    1785 ctgaagcctggcctgtgctctggggcgctgtcagactccggccggaactccatgtccagc 1844
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     387 CTGAAGCCTGGCCTGTGCTCTGGGGCGCTGTCAGACTCCGGCCGGAACTCCATGTCCAGC 328

Qy    1845 ctgcccacacacagcaccagcagcagctaccagctggacccgctggtcacacccgtggga 1904
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     327 CTGCCCACACACAGCACCAGCAGCAGCTACCAGCTGGACCCGCTGGTCACACCCGTGGGA 268

Qy    1905 cccacaagccgttttgggggctccgcccacaacatcacccagggcatcgtcctccaggac 1964
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     267 CCCACAAGCCGTTTTGGGGGCTCCGCCCACAACATCACCCAGGGCATCGTCCTCCAGGAC 208

Qy    1965 agcaacatgatgagcctgaaggctctgtccttctccgacggaggtagcaagctgggccac 2024
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     207 AGCAACATGATGAGCCTGAAGGCTCTGTCCTTCTCCGACGGAGGTAGCAAGCTGGGCCAC 148

Qy    2025 tcgaacaaggcagacaagggcccctcgtgtgtccgctcccccatctccacggacgagtgc 2084
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     147 TCGAACAAGGCAGACAAGGGCCCCTCGTGTGTCCGCTCCCCCATCTCCACGGACGAGTGC 88

Qy    2085 agcatccaggagctggaacagaagctgttggagagggagggcgccctccagaagctgcag 2144
           ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||
Db      87 AGCATCCAGGAGCTGGAGCAGAAGCTGTTGGAGAGGGAGGGCGCCCTCCAGAAGCTGCAG 28

Qy    2145 cgcagctttgaggagaaggagcttgcc 2171
           |||||||||||||||||||||||||||
Db      27 CGCAGCTTTGAGGAGAAGGAGCTTGCC 1
```

```
RESULT  14
AI360882/c
LOCUS       AI360882      442 bp    mRNA            EST       15-FEB-1999
DEFINITION  qy01c12.x1 NCI_CGAP_Brn23 Homo sapiens cDNA clone IMAGE:2010742 3',
            mRNA sequence.
ACCESSION   AI360882
VERSION     AI360882.1  GI:4112503
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 442)
  AUTHORS   NCI/NINDS-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute / National Institute of Neurological
            Disorders and Stroke, Brain Tumor Genome Anatomy Project
            (CGAP/BTGAP), Tumor Gene Index
  JOURNAL   Unpublished (1998)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert_Strausberg@nih.gov
            Tissue Procurement: David N. Louis, M.D., Myrna R. Rosenfeld M.D.,
            Ph.D.
             cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
            Bonaldo, Ph.D.
             cDNA Library Arrayed by: Greg Lennon, Ph.D.
             DNA Sequencing by: Washington University Genome Sequencing Center
             Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html
            Insert Length: 707   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 415.
FEATURES             Location/Qualifiers
     source          1. .442
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2010742"
                     /clone_lib="NCI_CGAP_Brn23"
                     /tissue_type="glioblastoma (pooled)"
                     /lab_host="DH10B"
                     /note="Organ: brain; Vector: pT7T3D-Pac (Pharmacia) with a
                     modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
                     strand cDNA was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCATATCTTTTTTTTTTTTTTTTTTTTT
                     T 3']; double-stranded cDNA was ligated to Eco RI
                     adaptors (Pharmacia), digested with Not I and cloned into
                     the Not I and Eco RI sites of the modified pT7T3 vector.
                     Library is normalized, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT      126 a    124 c    109 g     83 t
ORIGIN


  Query Match             4.9%;  Score 440.4;  DB 19;  Length 442;
  Best Local Similarity   99.8%;  Pred. No. 3e-35;
  Matches  441;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy    8602 aggcactcagtgagcctcctgtgcctggccccagtctgggccatctcttaggtgagacag 8661
           |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Db     442 AGGCACTCAGTGAGCCTCCTGTGCCTGGCCCCAGCCTGGGCCATCTCTTAGGTGAGACAG 383

Qy    8662 ttgcccgaaactaagccaggcctggctggaggagcagcagcttggggagagggatttccc 8721
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     382 TTGCCCGAAACTAAGCCAGGCCTGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCC 323

Qy    8722 tgcagacctcaagccatcatgcggtgggtgctgccatgacagaggctgcacccctgggcc 8781
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     322 TGCAGACCTCAAGCCATCATGCGGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCC 263

Qy    8782 agcggggctgctcacccacctcttgtgcaaggtggcctttgtgctgcgcctgcaggcaga 8841
```

```
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    262 AGCGGGGCTGCTCACCCACCTCTTGTGCAAGGTGGCCTTTGTGCTGCGCCTGCAGGCAGA 203

Qy   8842 gctggagcccccagcagaggcaggctgggacggaccagcatctggaagatgtacatagtt 8901
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    202 GCTGGAGCCCCCAGCAGAGGCAGGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTT 143

Qy   8902 atttttctctttgtggtttcttgtttggtttggtttgcttttgacagcttcattttattt 8961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    142 ATTTTTCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTT 83

Qy   8962 ttgacgtcacttttttggccatgtaaactatttgtggcaattttatgtttttatttatgaa 9021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     82 TTGACGTCACTTTTTGGCCATGTAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAA 23

Qy   9022 taaagaatgccatttctcacgc 9043
          |||||||||||||||||||||||
Db     22 TAAAGAATGCCATTTCTCACGC 1
```

RESULT 15
AI633596/c
LOCUS AI633596 453 bp mRNA EST 13-DEC-1999
DEFINITION th70a03.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:2123980 3', mRNA sequence.
ACCESSION AI633596
VERSION AI633596.1 GI:4684926
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 453)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 958 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 427.

FEATURES Location/Qualifiers
source 1. .453
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2123980"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NbHM, pregnant uterus NbHPU, and fetal heart NbHH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of I.M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479."

BASE COUNT 126 a 123 c 110 g 94 t
ORIGIN

Query Match 4.9%; Score 439.8; DB 23; Length 453;
Best Local Similarity 99.5%; Pred. No. 3.4e-35;
Matches 441; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
Qy   8602 aggcactcagtgagcctcctgtgcctggccccagtctgggccatctcttaggtgagacag 8661
          ||||||||||||||||||||||||||||||||||| ||||||||||||||||||||||||
Db    453 AGGCACTCAGTGAGCCTCCTGTGCCTGGCCCCAGCCTGGGCCATCTCTTAGGTGAGACAG 394

Qy   8662 ttgcccgaaactaagccaggcctggctggaggagcagcagcttggggagagggatttccc 8721
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    393 TTGCCCGAAACTAAGCCAGGCCTGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCC 334

Qy   8722 tgcagacctcaagccatcatgcggtgggtgctgccatgacagaggctgcacccctgggcc 8781
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    333 TGCAGACCTCAAGCCATCATGCGGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCC 274

Qy   8782 agcggggctgctcacccacctcttgtgcaaggtggcctttgtgctgcgcctgcaggcaga 8841
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    273 AGCGGGGCTGCTCACCCACCTCTTGTGCAAGGTGGCCTTTGTGCTGCGCCTGCAGGCAGA 214

Qy   8842 gctggagcccccagcagaggcaggctgggacggaccagcatctggaagatgtacatagtt 8901
          |||||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Db    213 GCTGGAGCCCCCAGCAAAGGCAGGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTT 154

Qy   8902 atttttctctttgtggtttcttgtttggtttggtttgcttttgacagcttcattttattt 8961
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    153 ATTTTTCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTT 94

Qy   8962 ttgacgtcacttttttggccatgtaaactatttgtggcaattttatgtttttatttatgaa 9021
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     93 TTGACGTCACTTTTTGGCCATGTAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAA 34

Qy   9022 taaagaatgccatttctcacgcc 9044
          |||||||||||||||||||||||
Db     33 TAAAGAATGCCATTTCTCACGCC 11
```

Search completed: April 5, 2001, 21:15:33
Job time: 6823 sec

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:52:25 ; Search time 318.41 Seconds
(without alignments)
10674.903 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*
1: /cgn2_2/gcgdata/geneseq/geneseqn/NA1980.DAT:*
2: /cgn2_2/gcgdata/geneseq/geneseqn/NA1981.DAT:*
3: /cgn2_2/gcgdata/geneseq/geneseqn/NA1982.DAT:*
4: /cgn2_2/gcgdata/geneseq/geneseqn/NA1983.DAT:*
5: /cgn2_2/gcgdata/geneseq/geneseqn/NA1984.DAT:*
6: /cgn2_2/gcgdata/geneseq/geneseqn/NA1985.DAT:*
7: /cgn2_2/gcgdata/geneseq/geneseqn/NA1986.DAT:*
8: /cgn2_2/gcgdata/geneseq/geneseqn/NA1987.DAT:*
9: /cgn2_2/gcgdata/geneseq/geneseqn/NA1988.DAT:*
10: /cgn2_2/gcgdata/geneseq/geneseqn/NA1989.DAT:*
11: /cgn2_2/gcgdata/geneseq/geneseqn/NA1990.DAT:*
12: /cgn2_2/gcgdata/geneseq/geneseqn/NA1991.DAT:*
13: /cgn2_2/gcgdata/geneseq/geneseqn/NA1992.DAT:*
14: /cgn2_2/gcgdata/geneseq/geneseqn/NA1993.DAT:*
15: /cgn2_2/gcgdata/geneseq/geneseqn/NA1994.DAT:*
16: /cgn2_2/gcgdata/geneseq/geneseqn/NA1995.DAT:*
17: /cgn2_2/gcgdata/geneseq/geneseqn/NA1996.DAT:*
18: /cgn2_2/gcgdata/geneseq/geneseqn/NA1997.DAT:*
19: /cgn2_2/gcgdata/geneseq/geneseqn/NA1998.DAT:*
20: /cgn2_2/gcgdata/geneseq/geneseqn/NA1999.DAT:*
21: /cgn2_2/gcgdata/geneseq/geneseqn/NA2000.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	365.8	4.0	393	16	T23583	Human gene signatu
c	2	212.2	2.3	700	20	X30339	DNA encoding a hum
c	3	211.2	2.3	54548	21	Z45596	DNA sequence of th
c	4	209	2.3	2116	20	X87623	Novel CD24* ORF en
c	5	209	2.3	2194	20	Z77555	Human ovarian tumo
c	6	209	2.3	2233	20	Z24431	Human bladder tumo
c	7	206.8	2.3	45546	20	X23520	Human kidney amino
	8	205.8	2.3	9365	21	Z50359	Human CD39-L4 geno
	9	205.4	2.3	138169	21	A34791	Human adenosine re
	10	205.4	2.3	141589	21	A35005	Human adenosine re
	11	205.4	2.3	141589	21	A35030	Human adenosine re
c	12	204.2	2.3	1601	21	A35191	Human adenosine re
c	13	204.2	2.3	1601	21	Z46814	Interleukin-10 (IL
c	14	204.2	2.3	1618	12	Q10207	pH15C insert conta
c	15	204.2	2.3	1618	14	Q46958	Human cytokine syn
c	16	204.2	2.3	1618	21	Z86905	Human CSIF coding
c	17	204.2	2.3	1645	21	Z58659	Human interleukin-
c	18	204.2	2.3	15630	21	A35195	Human adenosine re
	19	203.8	2.3	7042	20	V84785	Apoptosis inducer
	20	203.8	2.3	7042	21	Z46656	Human full-length
	21	203.8	2.3	7075	20	V84798	Apoptosis inducer
	22	203.6	2.3	16891	20	X37084	MEFV gene sequence
c	23	203.4	2.2	6235	21	Z29169	Human G-CSF genomi
c	24	203.4	2.2	6679	21	Z29170	Partial sequence o
c	25	203	2.2	65921	21	Z89046	Human nibrin DNA.
c	26	201.6	2.2	53526	19	T94101	Human PKD1 gene.
c	27	201.6	2.2	53577	17	T18551	Human polycystic k
c	28	201.6	2.2	53577	19	T94108	Human PKD1 locus b
c	29	200.8	2.2	1726	20	Z06248	Human secreted pro
c	30	200.6	2.2	781	20	Z15116	Human gene express
c	31	200.6	2.2	1296	19	V29031	Human protein comp
c	32	200.4	2.2	1189	20	Z00813	Human secreted pro
c	33	200	2.2	7849	16	Q94109	hML genomic DNA.
	34	199	2.2	283	15	Q63862	AP2 sequence obtd.
	35	199	2.2	22481	17	T11658	PEDF full length s
	36	199	2.2	162450	21	Z86967	Retinoblastoma bin
c	37	197.2	2.2	84607	20	X90847	Human PACAP genomi
c	38	196.2	2.2	730	21	Z97408	Human prostate can
c	39	196.2	2.2	745	21	Z97407	Human prostate can
c	40	196.2	2.2	4285	20	Z25338	Human nephrin nucl
c	41	196	2.2	2867	19	V68599	Lecithin-cholester
c	42	195.8	2.2	1600	20	V69205	Sequence ID No: 11
c	43	195.8	2.2	1600	21	Z24361	Human ICAM-R leade
c	44	195.8	2.2	22481	17	T11658	PEDF full length s
c	45	195.8	2.2	106746	21	A10225	Human PCTA-1 genom

ALIGNMENTS

RESULT 1
T23583
ID T23583 standard; cDNA to mRNA; 393 BP.
XX
AC T23583;
XX
DT 02-SEP-1996 (first entry)
XX
DE Human gene signature HUMGS05436.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN WO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues

XX
PS Claim 1; Page 1404; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 393 BP; 73 A; 117 C; 121 G; 78 T; 4 other;

Query Match 4.0%; Score 365.8; DB 16; Length 393;
Best Local Similarity 97.2%; Pred. No. 2.1e-56;
Matches 381; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy   8454 gatcggactctgcccctggacctgggaacgactggactgtcacggggttccctcctagct 8513
          ||||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 gatcggactctgcccctggacctgggaacgactggactgtcacggggttacctcctagct 60

Qy   8514 ctcccagtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggact 8573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctcccagtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggact 120

Qy   8574 gggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggcccc 8633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    121 gggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggnccc 180

Qy   8634 agtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggagg 8693
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggagg 240

Qy   8694 agcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgct 8753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgct 300

Qy   8754 gccatgacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaagg 8813
          |||||||||||||||| ||| |||||||||||||| |||||| |||||| ||||||||||
Db    301 gccatgacagaggctggaccnctgggccagcggggttgctcaaccacctttgtgcaagg 360

Qy   8814 tggcctttgtgc-tgcgcctgcaggcagagct 8844
          |||||||||||| |||||||||||||||  |
Db    361 tggcctttgtgcttgcgcctgcaggcagangt 392
```

RESULT 2
X30339/c
ID X30339 standard; DNA; 700 BP.
XX
AC X30339;
XX
DT 14-MAY-1999 (first entry)
XX
DE DNA encoding a human secreted protein.
XX
KW Secreted protein; cancer; tumour; neurodegenerative disorder;
KW developmental abnormality; foetal deficiency; blood disorder;
KW CNS disorder; immune system disease; autoimmune disease; hepatic disease;
KW renal disease; diabetes; inflammation; allergy; ischemic shock;
KW Alzheimer's; cognitive disorder; schizophrenia; cardiovascular disorder;
KW prostate disease; asthma; osteoporosis; arthritis; ss.
XX
OS Homo sapiens.

XX
PN WO9907891-A1.
XX
PD 18-FEB-1999.
XX
PF 04-AUG-1998; 98WO-US16235.
XX
PR 19-AUG-1997; 97US-0056732.
PR 05-AUG-1997; 97US-0054798.
PR 05-AUG-1997; 97US-0054803.
PR 05-AUG-1997; 97US-0054804.
PR 05-AUG-1997; 97US-0054806.
PR 05-AUG-1997; 97US-0054807.
PR 05-AUG-1997; 97US-0054808.
PR 05-AUG-1997; 97US-0054809.
PR 05-AUG-1997; 97US-0055309.
PR 05-AUG-1997; 97US-0055310.
PR 05-AUG-1997; 97US-0055312.
PR 05-AUG-1997; 97US-0055386.
PR 05-AUG-1997; 97US-0055311.
PR 18-AUG-1997; 97US-0055970.
PR 18-AUG-1997; 97US-0055986.
PR 19-AUG-1997; 97US-0056365.
PR 19-AUG-1997; 97US-0056366.
PR 19-AUG-1997; 97US-0056557.
PR 19-AUG-1997; 97US-0056370.
PR 19-AUG-1997; 97US-0056371.
PR 19-AUG-1997; 97US-0056563.
PR 19-AUG-1997; 97US-0056731.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Brewer LA, Ebner R, Ferrie AM, Greene JM, Janat F, Ni J;
PI Olsen HS, Rosen CA, Ruben SM, Soppet DR, Young PE, Yu G;
XX
DR WPI; 1999-167452/14.
DR P-PSDB; Y10818.
XX
PT New isolated human genes encoding secreted polypeptides - useful for
PT diagnosis and treatment of pathalogical diseases
XX
PS Claim 3; Page 242; 331pp; English.
XX
CC The specification describes secreted proteins and their corresponding
CC polynucleotides which are useful for preventing, treating or ameliorating
CC medical conditions, e.g. by protein or gene therapy. Pathological
CC conditions can also be diagnosed by determining the amount of the
CC secreted polypeptides in a sample or by determining the presence of
CC mutations in the polynucleotides. Specific uses are described for each
CC of the products, based on which tissues they are most highly
CC expressed in, and include developing products for the diagnosis or
CC treatment of cancer, tumours, neurodegenerative disorders, developmental
CC abnormalities and foetal deficiencies, blood disorders, CNS disorders,
CC diseases of the immune system, autoimmune diseases, hepatic and renal
CC disease, diabetes, inflammation, allergies, ischemic shock, Alzheimer's
CC and cognitive disorders, schizophrenia, cardiovascular disorders,
CC prostate diseases, asthma, disorders involving osteoclasts such as
CC osteoporosis, arthritis or malignancies, diseases of testes, lung or
CC thymus, digestive/endocrine disorders, infections and AIDS. The
CC polypeptides are also useful for identifying their binding partners.
XX
SQ Sequence 700 BP; 234 A; 146 C; 147 G; 173 T; 0 other;

Query Match 2.3%; Score 212.2; DB 20; Length 700;
Best Local Similarity 75.1%; Pred. No. 3.8e-29;
Matches 265; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```
Qy   6880 ttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtggca 6939
          || |||||  ||||  || |||| |||||||||| || ||||||||||||||||||||||
Db    692 TTTTTTTTTTTTTTGGAAACAGAGTCTTGCTCTATCACCCAGGCTGGAGTGCAGTGGCG 633

Qy   6940 tgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagcct 6999
```

```
          ||||| | |||||| ||| ||||   |||||||| ||| || || ||||||||||||||||
Db    632 TGATCTTGGCTCACCGCAACCTCCATCTCCTGGGTTCAGGCGATTCTCCTGCCTCAGCCT 573

Qy   7000 tccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttttt 7059
          || | || ||||||||||||   ||| ||||  |||||||||  ||||  | |||||
Db    572 CCCGAGTACCTGGGACTACAGGCATGCGTCACCACGCCTGGCTATTTTTTTGTATTTTTA 513

Qy   7060 gtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagcaat 7119
          |||| || ||| | ||    |||  | | ||||||||  ||||  || ||||| |||||
Db    512 GTAGAGATGGGTTTTCACCATGTTAGCCAGGCTGGTCTTGAACTCCTGACCTCAGGCAAT 453

Qy   7120 ccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggcca 7179
          || || ||||||| |||||||| ||||| |||| ||| ||||||||||||  || ||||
Db    452 CCTCCCGCCTTGGTCTCCCAAAATGCTGGGATTACAGGTGTGAGCCACCGCACCCGGCCG 393

Qy   7180 gatttttctttttattcttctttttcttttttcttttttgctttcttgtctttttcag 7232
          ||   |  || | | | ||| ||| ||| ||||||    |  ||| ||  | |
Db    392 GAAAGTAATTGTTTACATCTCTCTCTTTTTTTTTTTTAATTGTGTTTTGACTG 340


RESULT   3
Z45596/c
ID   Z45596 standard; DNA; 54548 BP.
XX
AC   Z45596;
XX
DT   06-APR-2000  (first entry)
XX
DE   DNA sequence of the wild type human PCTG4 region of Xq13.
XX
KW   Human; PCTG4 region; X chromosome; q13 region; polymorphism;
KW   mental retardation; autism; depression; bipolar affective disorder;
KW   hypothyroidism; OPA gene; neuropsychiatric disorder; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   exon            1391..1490
FT                   /*tag= a
FT                   /number= 1
FT                   /note= "human OPA exon"
FT   promoter        1398..1648
FT                   /*tag= b
FT                   /note= "human OPA promoter"
FT   promoter        complement (1565..1815)
FT                   /*tag= c
FT                   /note= "unknown promoter"
FT   exon            2008..2112
FT                   /*tag= d
FT                   /number= 2
FT                   /note= "human OPA promoter"
FT   exon            2321..2512
FT                   /*tag= e
FT                   /number= 3
FT                   /note= "human OPA promoter"
FT   exon            2649..2805
FT                   /*tag= f
FT                   /number= 4
FT                   /note= "human OPA promoter"
FT   exon            3608..3787
FT                   /*tag= g
FT                   /number= 5
FT                   /note= "human OPA promoter"
FT   exon            3962..4072
FT                   /*tag= h
FT                   /number= 6
FT                   /note= "human OPA promoter"
FT   exon            4197..4451
FT                   /*tag= i
FT                   /number= 7
FT                   /note= "human OPA promoter"
FT   exon            4839..4986
FT                   /*tag= j
FT                   /number= 8
FT                   /note= "human OPA promoter"
FT   exon            5148..5247
FT                   /*tag= k
FT                   /number= 9
FT                   /note= "human OPA promoter"
FT   exon            5377..5513
FT                   /*tag= l
FT                   /number= 10
FT                   /note= "human OPA promoter"
FT   exon            5734..5865
FT                   /*tag= m
FT                   /number= 11
FT                   /note= "human OPA promoter"
FT   exon            6233..6359
FT                   /*tag= n
FT                   /number= 12
FT                   /note= "human OPA promoter"
FT   exon            6798..7027
FT                   /*tag= o
FT                   /number= 13
FT                   /note= "human OPA promoter"
FT   exon            7403..7483
FT                   /*tag= p
FT                   /number= 14
FT                   /note= "human OPA promoter"
FT   exon            7615..7784
FT                   /*tag= q
FT                   /number= 15
FT                   /note= "human OPA promoter"
FT   exon            7989..8133
FT                   /*tag= r
FT                   /number= 16
FT                   /note= "human OPA promoter"
FT   exon            8301..8351
FT                   /*tag= s
FT                   /number= 17
FT                   /note= "human OPA promoter"
FT   exon            8674..8792
FT                   /*tag= t
FT                   /number= 18
FT                   /note= "human OPA promoter"
FT   exon            8979..9122
FT                   /*tag= u
FT                   /number= 19
FT                   /note= "human OPA promoter"
FT   exon            9606..9769
FT                   /*tag= v
FT                   /number= 20
FT                   /note= "human OPA promoter"
FT   exon            9973..10104
FT                   /*tag= w
FT                   /number= 21
FT                   /note= "human OPA promoter"
FT   exon            10530..10757
FT                   /*tag= x
FT                   /number= 22
FT                   /note= "human OPA promoter"
FT   exon            10933..11077
FT                   /*tag= y
FT                   /number= 23
FT                   /note= "human OPA promoter"
FT   exon            11235..11355
FT                   /*tag= z
FT                   /number= 24
FT                   /note= "human OPA promoter"
FT   exon            11751..12065
FT                   /*tag= aa
FT                   /number= 25
FT                   /note= "human OPA promoter"
FT   exon            12317..12492
FT                   /*tag= ab
FT                   /number= 26
```

```
FT                   /note= "human OPA promoter"
FT   exon            12672..12851
FT                   /*tag= ac
FT                   /number= 27
FT                   /note= "human OPA promoter"
FT   exon            14187..14258
FT                   /*tag= ad
FT                   /number= 28
FT                   /note= "human OPA promoter"
FT   exon            14710..14843
FT                   /*tag= ae
FT                   /number= 29
FT                   /note= "human OPA promoter"
FT   exon            15014..15175
FT                   /*tag= af
FT                   /number= 30
FT                   /note= "human OPA promoter"
FT   exon            15487..15598
FT                   /*tag= ag
FT                   /number= 31
FT                   /note= "human OPA promoter"
FT   exon            15765..15854
FT                   /*tag= ah
FT                   /number= 32
FT                   /note= "human OPA promoter"
FT   exon            16999..17109
FT                   /*tag= ai
FT                   /number= 33
FT                   /note= "human OPA promoter"
FT   exon            17351..17492
FT                   /*tag= aj
FT                   /number= 34
FT                   /note= "human OPA promoter"
FT   exon            17736..17897
FT                   /*tag= ak
FT                   /number= 35
FT                   /note= "human OPA promoter"
FT   exon            18927..19301
FT                   /*tag= al
FT                   /number= 36
FT                   /note= "human OPA promoter"
FT   exon            19525..19675
FT                   /*tag= am
FT                   /number= 37
FT                   /note= "human OPA promoter"
FT   exon            19833..20029
FT                   /*tag= an
FT                   /number= 38
FT                   /note= "human OPA promoter"
FT   exon            20207..20281
FT                   /*tag= ao
FT                   /number= 39
FT                   /note= "human OPA promoter"
FT   exon            20372..20590
FT                   /*tag= ap
FT                   /number= 40
FT                   /note= "human OPA promoter"
FT   exon            23283..23505
FT                   /*tag= aq
FT                   /number= 41
FT                   /note= "human OPA promoter"
FT   exon            23878..24018
FT                   /*tag= ar
FT                   /number= 42
FT                   /note= "human OPA promoter"
FT   exon            24531..24612
FT                   /*tag= as
FT                   /number= 43
FT                   /note= "human OPA promoter"
FT   exon            24823..25096
FT                   /*tag= at
FT                   /number= 44
FT                   /note= "human OPA promoter"
FT   promoter        26123..26275
FT                   /*tag= au
FT                   /note= "probable NL-3 promoter"

  Query Match             2.3%;  Score 211.2;  DB 21;  Length 54548;
  Best Local Similarity   78.8%;  Pred. No. 1.3e-28;
  Matches  279;  Conservative    0;  Mismatches   68;  Indels    7;  Gaps    2;

Qy    4317 ggctaggcgtggtggctcacgcctgcaatcccagcactttgggaggcagaggcaggcaga 4376
           ||| |||||||||||||| ||||| ||||||||||||||||||||| ||||||||||||
Db   39536 GGCCAGGCGTGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGGCAGGCAGA 39477

Qy    4377 gcacttgaggtcaggaattcgagatcagcctggccaacatgaggaaaccccgtctctact 4436
           ||| | |||||||||| ||||||| |||||||||||||||  ||||||| |||||||||
Db   39476 TCACCTAAGGTCAGGAGTTCGAGACCAGCCTGGCCAACATGGTGAAACCCTGTCTCTACT 39417

Qy    4437 aaaaatacaaaaatcagccaggcatggtggcaggtgtctgtaatcccagctacttgggag 4496
           |||||||||||||| |||| |||||||| ||||||| |||||||||||||||||  ||||
Db   39416 AAAAATACAAAAATTAGCCTGGCATGGTGAGCAGGTGCCTGTAATCCCAGCTACTCAGGAG 39357

Qy    4497 gctgaggcaggagaatcag------aggggaggcggaggttgcagtgagccaagatcacg 4550
           ||||||||||||||||||        |||||||||||||||||||||||||||||   |
Db   39356 GCTGAGGCAGGAGAATCAGTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCCAAGATGGTG 39297

Qy    4551 ccactacaccccagcctaggtgacaaagcgagact-tctcaaatattaacaataataata 4609
           ||| | ||| |||||| || |||| || |||||| |  |  | |   | ||||| ||||
Db   39296 CCATTGCACTCCAGCCTGGGCGACAGAGTGAGACTCTGCCTCAAAAACAAAATAAAAATA 39237

Qy    4610 tactatgtgtcattatacatgatgattattattttatcattttactatatagcc 4663
             ||   | | || | |     || | ||  | | ||  |||  |||  |||
Db   39236 ATAAATAAATAAATAAATAAATAAATAAATAAATAAGCAAATTAAAATAAGGCC 39183


RESULT   4
X87623/c
ID   X87623 standard; DNA; 2116 BP.
XX
AC   X87623;
XX
DT   26-OCT-1999  (first entry)
XX
DE   Novel CD24* ORF encoding retinoblastoma interacting protein.
XX
KW   CD24*; retinoblastoma interacting protein; RB-IP; human;
KW   cell cycle; cell differentiation; signal transduction;
KW   cell proliferation; psoriasis; malignancy; cancer;
KW   neurodegeneration; viral infection; therapy; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             57..299
FT                   /*tag= a
FT                   /product= CD24
FT   CDS             957..1151
FT                   /*tag= b
FT                   /product= CD24*
FT                   /note= "codes for Y06601, an isolated nucleic acid
FT                          comprising this nucleotide sequence is
FT                          specifically claimed in Claim 30"
XX
PN   WO9941376-A2.
XX
PD   19-AUG-1999.
XX
PF   12-FEB-1999;   99WO-US03072.
XX
PR   11-FEB-1999;   99US-0074559.
PR   12-FEB-1998;   98US-0074559.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Nandabalan K,  Schulz VP,  Yang M;
```

XX
DR WPI; 1999-494537/41.
DR P-PSDB; Y06601.
XX
PT New complexes of retinoblastoma protein with its binding proteins
PT and related antibodies, nucleic acids and modulators, used for
PT diagnosis and treatment of cancer, neurodegeneration and viral
PT infection
XX
PS Claim 30; Fig 1; 103pp; English.
XX
CC This is the nucleotide sequence of the human CD24 gene, in which a
CC novel open reading frame has been identified that encodes a protein
CC (see Y06601) designated CD24*. CD24* was initially identified as a
CC retinoblastoma interacting protein (RB-IP) using an improved,
CC modified yeast two-hybrid assay system. The invention relates to
CC complexes of RB and RB-IPs including CD24*. The complexes modulate
CC function of RB, i.e. cell-cycle control, cell differentiation and
CC apoptosis, resistance to viral infection, intracellular signal
CC transduction, transcriptional control, tumorigenesis and
CC degeneration. Detecting abnormal levels, or function, of the
CC RB-RB-IP complex, or levels of related RNA, is used to diagnose or
CC monitor treatment of (pre) cancer or other hyperproliferative states
CC such as psoriasis, neurodegeneration and viral infection. These
CC conditions are also diagnosed by detecting levels of CD24*, Set*,
CC GluT1* or 115392* (see Y06601-04). These disorders may be treated
CC or prevented by administering an agent that modulates function of
CC the RB-RB-IP complex. Animals which contain inactivated RB or
CC RB-IP genes, or have these genes under control of non-native
CC promoters, are used to screen for modulators of the complex.
XX
SQ Sequence 2116 BP; 616 A; 448 C; 414 G; 638 T; 0 other;

Query Match 2.3%; Score 209; DB 20; Length 2116;
Best Local Similarity 73.6%; Pred. No. 1.7e-28;
Matches 282; Conservative 0; Mismatches 95; Indels 6; Gaps 1;

Qy 4275 aattattagtgaaagtatgtttataataatatacttctgtgtggctaggcgtggtggctc 4334
Db 1470 AAAAAGTATTCAAACATATGCAGAAATAAAAAGCATTTTGATGGCTGGGCGCGGGGGCTC 1411

Qy 4335 acgcctgcaatcccagcactttgggaggcagaggcaggcagagcacttgaggtcaggaat 4394
Db 1410 AAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGGTCAGGAGT 1351

Qy 4395 tcgagatcagcctggccaacatgaggaaaccccgtctctactaaaaatacaaaaatcagc 4454
Db 1350 TCTAAACCAGCCTGGGCAACATGGTGAAACCCTGTCTCTACTAAAATTACAAAAATTAGC 1291

Qy 4455 caggcatggtggcaggtgtctgtaatcccagctacttgggaggctgaggcaggagaatc- 4513
Db 1290 CGGGCATGGTGGCAGGTGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG 1231

Qy 4514 -----agaggggaggcggaggttgcagtgagccaagatcacgccactacaccccagccta 4568
Db 1230 CTTGAGCCCGGGAGGCAAATGGTACACTGATCTGAGATCGCACCACTGCACTCCGGCCTG 1171

Qy 4569 ggtgacaaagcgagacttctcaaatattaacaataataatatactatgtgtcattataca 4628
Db 1170 GGCGACAAAGTGAGACTGTCTAAAAAATAATAATAATAAAAAGCATTTTGAAATTAGTCG 1111

Qy 4629 tgatgattattattttatcattt 4651
Db 1110 TGGTCAATGCAATTCTACTCTTT 1088

RESULT 5
Z77555/c
ID Z77555 standard; cDNA; 2194 BP.
XX
AC Z77555;
XX
DT 10-APR-2000 (first entry)
XX
DE Human ovarian tumor cDNA library derived EST fragment 106.
XX
KW Expressed sequence tag; EST; human; ovarian tumor; anticancer;
KW gene therapy; treatment; ss.
XX
OS Homo sapiens.
XX
PN DE19817557-A1.
XX
PD 21-OCT-1999.
XX
PF 09-APR-1998; 98DE-1017557.
XX
PR 09-APR-1998; 98DE-1017557.
XX
PA (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;
XX
DR WPI; 1999-591920/51.
XX
PT New nucleic acid sequences expressed in ovarian, and some other, cancer
PT tissues, and derived polypeptides, for treatment of ovarian cancer and
PT identification of therapeutic agents -
XX
PS Claim 3; Page 228-229; 310pp; German.
XX
CC This invention describes novel nucleic acid (cDNA) sequences (A) which
CC have anticancer activity and are highly expressed in ovarian tumor
CC tissue (and some also in testis and breast cancer tissue). The products
CC of the invention can be used for gene therapy. (A) are used (i) for
CC recombinant expression of polypeptides (B) and (ii) to isolate complete
CC genes. (B) are used (i) to identify agents suitable for treatment of
CC ovarian cancer; (ii) directly for treating this form of cancer (including
CC expression from gene therapy vectors) and (iii) for generation of
CC specific antibodies. (A) are identified by assembling ESTs (expressed
CC sequence tags) from a particular tissue type before comparison of
CC expression patterns. This allows a significantly longer fragment of the
CC gene to be revealed, so should reduce the number of failures associated
CC with the fact that ESTs from different libraries may represent different
CC parts of the same unknown gene, distorting the estimated frequency of
CC occurrence in a particular tissue. Z77450-Z77572 represent the human
CC ovarian tumor cDNA library derived EST fragments described in the method
CC of the invention and encode the protein fragments represented in
CC Y76505-Y76638.
XX
SQ Sequence 2194 BP; 686 A; 431 C; 416 G; 661 T; 0 other;

Query Match 2.3%; Score 209; DB 20; Length 2194;
Best Local Similarity 76.0%; Pred. No. 1.7e-28;
Matches 273; Conservative 0; Mismatches 80; Indels 6; Gaps 1;

Qy 4299 aataatatacttctgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgg 4358
Db 1463 AATAAAAAGCATTTTGATGGCTGGGCGCGGGGGCTCAAGCCTGTAATCCCAGCACTTTGG 1404

Qy 4359 gaggcagaggcaggcagagcacttgaggtcaggaattcgagatcagcctggccaacatga 4418
Db 1403 GAGGCCGAGGCGGGTGGATCACCTGAGGTCAGGAGTTCGAAACCAGCCTGGGCAACATGG 1344

Qy 4419 ggaaaccccgtctctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgta 4478
Db 1343 TGAAACCCTGTCTCTACTAAAATTACAAAAATTAGCCGGGCATGGTGGCAGGTGCCTGTA 1284

Qy 4479 atcccagctacttgggaggctgaggcaggagaatc------agaggggaggcggaggttg 4532
Db 1283 ATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCGGGAGGCAAATGGTA 1224

Qy 4533 cagtgagccaagatcacgccactacaccccagcctaggtgacaaagcgagacttctcaaa 4592

```
Db   1223 CACTGATCTGAGATCGCACCACTGCACTCCGGCCTGGGCGACAAAGTGAGACTGTCTAAA 1164

Qy   4593 tattaacaataataatatactatgtgtcattatacatgatgattattattttatcattt 4651
          | ||| |||||||| |  |  | ||  ||||  |  | | | |   ||| ||   |||
Db   1163 AAATAATAATAATAAAAAGCATTTTGAAATTAGTCGCGGTCAATGCAATTCTACTCTTT 1105


RESULT   6
Z24431/c
ID   Z24431 standard; cDNA; 2233 BP.
XX
AC   Z24431;
XX
DT   14-FEB-2000  (first entry)
XX
DE   Human bladder tumour cDNA library derived EST 43.
XX
KW   Expressed sequence tag; human; bladder; tumour; cancer; cytostatic;
KW   treatment; gene therapy; EST; ss.
XX
OS   Homo sapiens.
XX
PN   DE19818619-A1.
XX
PD   28-OCT-1999.
XX
PF   21-APR-1998;   98DE-1018619.
XX
PR   21-APR-1998;   98DE-1018619.
XX
PA   (META-) METAGEN GES GENOMFORSCHUNG MBH.
XX
PI   Rosenthal A,  Specht T,  Hinzmann B,  Schmitt A,  Pilarsky C,  Dahl E;
XX
DR   WPI; 1999-612028/53.
XX
PT   New nucleic acid sequences expressed in bladder tumor tissue, and
PT   derived polypeptides, for treatment of bladder tumor and identification
PT   of therapeutic agents  -
XX
PS   Claim 3; Page 93; 132pp; German.
XX
CC   This invention describes novel polypeptide fragments (I) and the
CC   polynucleotides (II) that encode them that are highly expressed in a
CC   human bladder tumour and which have cytostatic activity. (II) are used
CC   for recombinant expression of (I) and to isolate complete genes. (I) are
CC   used to identify agents suitable for treatment of bladder cancer, to
CC   directly treat this form of cancer (including expression from gene
CC   therapy vectors) or are used in a preparation for cancer treatment. (I)
CC   is also used for the generation of specific antibodies. (II) are
CC   identified by assembling ESTs (expressed sequence tags) from a particular
CC   tissue type before comparison of expression patterns. This allows a
CC   significantly longer fragment of the gene to be revealed, and therefore
CC   reduces the number of failures associated with the fact that ESTs from
CC   different libraries may represent different parts of the same unknown
CC   gene, distorting the estimated frequency of occurrence in a particular
CC   tissue. Z43260-Z43309 represent expressed sequence tag (EST) fragments
CC   isolated from a human bladder tumour cDNA library which encode the
CC   proteins represented in Y66143-Y66198.
XX
SQ   Sequence 2233 BP; 685 A; 431 C; 455 G; 662 T; 0 other;


  Query Match             2.3%;  Score 209;  DB 20;  Length 2233;
  Best Local Similarity   76.0%;  Pred. No. 1.7e-28;
  Matches  273;  Conservative    0;  Mismatches   80;  Indels    6;  Gaps    1;

Qy   4299 aataatatacttctgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgg 4358
          |||||  |  | | |   ||||| |||| || |||||| ||||| ||||||||||||||||
Db   1463 AATAAAAAGCATTTTGATGGCTGGGCGCGGGGGCTCAAGCCTGTAATCCCAGCACTTTGG 1404

Qy   4359 gaggcagaggcaggcagagcacttgaggtcaggaattcgagatcagcctggccaacatga 4418
          ||||| ||||| ||  || ||| ||||||||||| ||||| | |||||||| |||||||

Db   1403 GAGGCCGAGGCGGGTGGATCACCTGAGGTCAGGAGTTCGAAACCAGCCTGGGCAACATGG 1344

Qy   4419 ggaaaccccgtctctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgta 4478
          ||||||| |||||||||||||| ||||||||| |||| ||||||||||||||| |||||
Db   1343 TGAAACCCTGTCTCTACTAAAATTACAAAAATTAGCCGGGCATGGTGGCAGGTGCCTGTA 1284

Qy   4479 atcccagctacttgggaggctgaggcaggagaatc------agaggggaggcggaggttg 4532
          ||||||||||||||||||||||||||||||||||          |||||||  | | |
Db   1283 ATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCGCTTGAGCCCGGGAGGCAAATGGTA 1224

Qy   4533 cagtgagccaagatcacgccactacaccccagcctaggtgacaaagcgagacttctcaaa 4592
          || ||| |  ||||| | ||||| ||| || |||| || ||||||| ||||||   |||
Db   1223 CACTGATCTGAGATCGCACCACTGCACTCCGGCCTGGGCGACAAAGTGAGACTGTCTAAA 1164

Qy   4593 tattaacaataataatatactatgtgtcattatacatgatgattattattttatcattt 4651
          | ||| |||||||| |  |  | ||  ||||  |  | | | |   ||| ||   |||
Db   1163 AAATAATAATAATAAAAAGCATTTTGAAATTAGTCGCGGTCAATGCAATTCTACTCTTT 1105


RESULT   7
X23520/c
ID   X23520 standard; DNA; 45546 BP.
XX
AC   X23520;
XX
DT   23-JUN-1999  (first entry)
XX
DE   Human kidney aminopeptidase P genomic DNA fragment 4.
XX
KW   Aminopeptidase; human; AmP; gene therapy; treatment; AmP-deficiency;
KW   prenatal diagnosis; angioedema; antihypertensive agent; atherosclerosis;
KW   arterial stenosis; industrial protein feed; malabsorption syndrome;
KW   proteinaceous waste degradation; additive; immunohistochemistry; ss.
XX
OS   Homo sapiens.
XX
PN   WO9911799-A2.
XX
PD   11-MAR-1999.
XX
PF   02-SEP-1998;   98WO-US18426.
XX
PR   02-SEP-1997;   97US-0057854.
XX
PA   (MEDI-) MEDICAL COLLEGE GEORGIA RES INST.
XX
PI   Ryan JW,  Sprinkle TJC,  Venema RC;
XX
DR   WPI; 1999-205193/17.
XX
PT   Nucleic acid encoding human aminopeptidase P
XX
PS   Claim 13; Page 165-192; 201pp; English.
XX
CC   This invention describes the isolation of a novel human aminopeptidase P
CC   (AmP). This protein is used to produce recombinant AmP and can be used
CC   for gene therapy for treating AmP-deficiency conditions. Its fragments
CC   are used as primers and probes to identify patients with homozygous and
CC   heterozygous AmP deficiency, including prenatal diagnosis (patients
CC   defective in AmP are at risk of developing angioedema if treated with
CC   angiotensin-converting enzyme inhibitors), also as antisense inhibitors
CC   in cases of excessive AmP expression. The product of the invention is
CC   also used to identify AmP-expressing sequences in other animals and to
CC   generate transgenic animals, and comparisons of genomic sequences are
CC   used to detect mutations. AmP inhibitors are potentially useful as
CC   antihypertensive agents and to prevent or treat arterial (re)stenosis
CC   or atherosclerosis. The structure of AmP is used to design synthetic
CC   substrates, e.g. for use in AmP assays. AmP, which hydrolyzes N-terminal
CC   imido bonds, can be used to degrade industrial protein feeds to free
CC   amino acids, to degrade proteinaceous wastes, as additives in enzyme
CC   formulations used to treat malabsorption syndrome and for studying its
CC   biological role. Antibodies against AmP are used in immunohistochemical
CC   methods to study AmP distribution.
```

```
XX
SQ   Sequence 45546 BP; 12027 A; 11359 C; 11380 G; 10780 T; 0 other;


  Query Match             2.3%;  Score 206.8;  DB 20;  Length 45546;
  Best Local Similarity   70.3%;  Pred. No. 7.5e-28;
  Matches  305;  Conservative    0;  Mismatches  127;  Indels    2;  Gaps    2;

Qy   6855 ttctctgttgtttggttcagggattttatttttaattttatgagacagggtcttgctctg 6914
           ||  || || |||||||  ||  || | || ||   ||| |||||||| ||||||||||
Db   3731 TTGGCTCTTCTTTGGTTTTGGTTTTGTTTTGTTTTGTTTTTGAGACAGAGTCTTGCTCTG 3672

Qy   6915 tcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggc 6974
           || ||||| ||||||||||||||||  ||| | ||||||||| ||||    |||||||
Db   3671 TCGCCCAGACTGGAGTGCAGTGGCACAATCTTGGCTCACTGCAACCTCTGCTTCCTGGGT 3612

Qy   6975 tcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgt 7034
           ||||||||| || |||||| ||||| |  | |||| |||| ||||||  | |||||||
Db   3611 TCAAGCAATTCTTCTGCCTTAGCCTCCTGAGTAGCGGGGATTACAGGCACCCGCCACCAT 3552

Qy   7035 gcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctgg 7094
           |||||||||| ||||  | ||||| |||| |||||||| ||    ||||||| | |||||
Db   3551 GCCTGGCTAA-TTTTTTTATTTTTAGTAGAGACGGGGTTTCACCATGTTGCCCAGGCTGG 3493

Qy   7095 tctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgc 7154
           |||| ||||  || ||||| |  |||||||| || || ||||||||||||| |||| |
Db   3492 TCTCGAACTCCTGACCTCAGGTGATCCACCTGGCTCGGTCTCCCAAAGTGCTGGGATTAC 3433

Qy   7155 agatgtgagccaccgtgcctggccagatttttctttttat-tcttctttcttttttcttttt 7213
           ||  || ||||||||||  ||||    ||  || || |  | ||| ||||  ||  | |
Db   3432 AGGCGTCAGCCACCGTGCTCGGCCTCCGTTGGCTCTTCTAACCTCTGTCTTCCTCCCTCT 3373

Qy   7214 tgctttcttgtcttttcagaagcaagccagacctagcaggctgttccatgttctatttttt 7273
           | | ||||  ||| ||       | ||  |      | ||    | | || |    |
Db   3372 TCCCTTCTGTGCTTCTCTTTTTGCTGTCACTCTAGAATTTGTCTTTTCTTTCCTCTGCCT 3313

Qy   7274 gactgtagccacag 7287
           | |   ||||| ||
Db   3312 GGCCTCAGCCAAAG 3299


RESULT   8
Z50359
ID   Z50359 standard; DNA; 9365 BP.
XX
AC   Z50359;
XX
DT   18-MAY-2000  (first entry)
XX
DE   Human CD39-L4 genomic DNA.
XX
KW   CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase;
KW   ATP Diphosphohydrolase; ATPDase; adenosine diphosphate; ADP; treatment;
KW   platelet aggregation; antithrombotic; thrombosis; myocardial infarction;
KW   cerebral ischaemia; angina; vascular graft; extracorporeal circulation;
KW   molecular weight marker; nutritional supplement; tumour; prevention;
KW   drug targeting; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             72..8642
FT                   /*tag= a
FT                   /product= "Human CD39-L4 protein"
FT                   /note= "Coding region is interrupted with introns"
FT   exon            1..288
FT                   /*tag= b
FT                   /number= 1
FT   intron          289..1280
FT                   /*tag= c
FT                   /number= 1
FT   exon            1281..1580
FT                   /*tag= d
FT                   /number= 2
FT   intron          1581..1819
FT                   /*tag= e
FT                   /number= 2
FT   exon            1820..1855
FT                   /*tag= f
FT                   /number= 3
FT   intron          1856..2466
FT                   /*tag= g
FT                   /number= 3
FT   exon            2467..2555
FT                   /*tag= h
FT                   /number= 4
FT   intron          2556..2862
FT                   /*tag= i
FT                   /number= 4
FT   exon            2863..2942
FT                   /*tag= j
FT                   /number= 5
FT   intron          2943..3888
FT                   /*tag= k
FT                   /number= 5
FT   exon            3889..3950
FT                   /*tag= l
FT                   /number= 6
FT   intron          3951..4893
FT                   /*tag= m
FT                   /number= 6
FT   exon            4894..4995
FT                   /*tag= n
FT                   /number= 7
FT   intron          4996..5846
FT                   /*tag= o
FT                   /number= 7
FT   exon            5847..5987
FT                   /*tag= p
FT                   /number= 8
FT   intron          5988..6965
FT                   /*tag= q
FT                   /number= 8
FT   exon            6966..7138
FT                   /*tag= r
FT                   /number= 9
FT   intron          7139..8555
FT                   /*tag= s
FT                   /number= 9
FT   exon            8556..9365
FT                   /*tag= t
FT                   /number= 10
XX
PN   WO200004041-A2.
XX
PD   27-JAN-2000.
XX
PF   16-JUL-1999;   99WO-US16180.
XX
PR   16-JUL-1998;   98US-0118205.
PR   24-JUL-1998;   98US-0122449.
PR   04-FEB-1999;   99US-0244444.
PR   19-MAR-1999;   99US-0273447.
PR   09-JUL-1999;   99US-0350836.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Ford J,  Mulero J;
XX
DR   WPI; 2000-182397/16.
DR   P-PSDB; Y44849.
XX
PT   New nucleic acid encoding human CD39-like protein, useful for treating
PT   and preventing thrombotic disease -
XX
PS   Example 11; Page 112-119; 125pp; English.
```

```
XX
CC   The present sequence is the genomic DNA encoding CD39-L4 protein, an
CC   apyrase and/or nucleotide diphosphatase (NDPase). It is isolated from
CC   the human CITB BAC genomic library. It is a soluble ATP
CC   Diphosphohydrolases (ATPDase) and is involved in the hydrolysis of
CC   adenosine diphosphate (ADP), the agonist that causes platelet
CC   aggregation. CD39-L4 protein has 30% and 80% homology to human and
CC   murine CD39. It has platelet aggregation inhibition and antithrombotic
CC   activity. CD39-L4 is used to treat or prevent thrombosis, myocardial
CC   infarction, cerebral ischaemia and angina. It is also used in vitro, to
CC   maintain vascular grafts or during extracorporeal circulation, to
CC   hydrolyse NDP, as molecular weight markers and as nutritional
CC   supplements. It is used to identify therapeutic agents that bind and
CC   modulate CD39-L4. It is coupled to toxins for targeting drugs to tumours
CC   or other cells that express CD39-L4.
XX
SQ   Sequence 9365 BP; 2439 A; 2005 C; 2191 G; 2632 T; 98 other;


  Query Match             2.3%;  Score 205.8;  DB 21;  Length 9365;
  Best Local Similarity   71.4%;  Pred. No. 8.4e-28;
  Matches  245;  Conservative   34;  Mismatches   62;  Indels    2;  Gaps    2;

Qy   6879 tttatttttaatttttatgagacagggtcttgctc-tgtcccccaggctggagtgcagtgg 6937
Db   7427 tttattattatttttttgagacggagtctygctcttgttkccyrggctggagtgcartgg 7486

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
Db   7487 crygatcwcrgctcactgcarcctycrcctcctgggttcaagcgattctyctgcctcagc 7546

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaa-cttttcatttttt 7056
Db   7547 ctcccragtagctgggattacaggcrygcgccaccacgccyggctaatttttttgtattt 7606

Qy   7057 tttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagc 7116
Db   7607 ttagtagagacggggtttcaccatgttggccaggctrktctyraactyctgayctcaggt 7666

Qy   7117 aatccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctgg 7176
Db   7667 gatccacccrcctcrgcctcccaaagtgctrgrattayaggyrtgagccaccackccyrg 7726

Qy   7177 ccagatttttctttttattcttctttctttttctttttgcttt 7219
Db   7727 cctttttgktcgkttctttttttttchttttttttttttttt 7769


RESULT   9
A34791
ID   A34791 standard; DNA; 138169 BP.
XX
AC   A34791;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2480.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17712.
XX
PR   03-AUG-1998;   98US-0095212.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstriction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers -
XX
PS   Disclosure; Page 613-645; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. A32313 to A35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC   185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC   differ from the previously named sequences. SEQ ID NO:11 to 1680 (A32323
CC   to A33992) are specifically claimed ONs from the present invention.
CC   N.B. Sequences given in the disclosure of the present invention do not
CC   match up with their corresponding SEQ ID NO: sequences given in the
CC   sequence listing.
XX
SQ   Sequence 138169 BP; 41766 A; 28262 C; 26273 G; 41868 T; 0 other;


  Query Match             2.3%;  Score 205.4;  DB 21;  Length 138169;
  Best Local Similarity   74.9%;  Pred. No. 1.6e-27;
  Matches  257;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy   6855 ttctctgttgtttggttcagggattttatttttaattttatgagacagggtcttgctctg 6914
Db  47893 ttgtttgtttgtttatttgtttctttttttttttttttttttgagacagggtctcactctg 47952

Qy   6915 tcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggc 6974
Db  47953 ttgcccaggctggagtgcagtggtatgatctcggctcactgcagcctccatctccctggt 48012

Qy   6975 tcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgt 7034
Db  48013 tcaagtgattctcatgtctcagcctcccgagtagctgggactacaggtgcatgccactac 48072

Qy   7035 gcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctgg 7094
Db  48073 gcctggctaagattttgtatttttattagagaaggggttttgccatgttggccaagctgg 48132

Qy   7095 tctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgc 7154
Db  48133 tctcgaactcctaacctcaggtgatccacctgcctcagcctcccaaagtgctgagattat 48192

Qy   7155 agatgtgagccaccgtgcctggccagatttttctttttattctt 7197
Db  48193 aggcatgagccactatgccccacctggtttctccttatttatt 48235
```

```
RESULT 10
A35005
ID   A35005 standard; DNA; 141589 BP.
XX
AC   A35005;
XX
DT   28-JUL-2000 (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2694.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17712.
XX
PR   03-AUG-1998;   98US-0095212.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstriction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers -
XX
PS   Disclosure; Page 916-948; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. A32313 to A35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC   185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC   differ from the previously named sequences. SEQ ID NO:11 to 1680 (A32323
CC   to A33992) are specifically claimed ONs from the present invention.
CC   N.B. Sequences given in the disclosure of the present invention do not
CC   match up with their corresponding SEQ ID NO: sequences given in the
CC   sequence listing.
XX
SQ   Sequence 141589 BP; 42856 A; 28939 C; 26863 G; 42931 T; 0 other;

  Query Match          2.3%;  Score 205.4;  DB 21;  Length 141589;
  Best Local Similarity  74.9%;  Pred. No. 1.7e-27;
  Matches  257;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy    6855 ttctctgttgtttggttcagggatttttatttttaattttatgagacagggtcttgctctg 6914
           || | ||||  ||  ||      |||| |||||  |||| ||||||||||||  |||||
Db   47893 ttgtttgtttgtttatttgtttctttttttttttttttttttgagacagggtctcactctg 47952

Qy    6915 tcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggc 6974
           |  |||||||||||||||||||| ||||||   ||||||||||||||   ||||  ||
Db   47953 ttgcccaggctggagtgcagtggtatgatctcggctcactgcagcctccatctccctggt 48012

Qy    6975 tcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgt 7034
           |||||  || ||| || |||||||| || | ||||||||||||||||||  |||||
Db   48013 tcaagtgattctcatgtctcagcctcccgagtagctgggactacaggtgcatgccactac 48072

Qy    7035 gcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctgg 7094
           ||||||||||  |||  | |||||  ||| || ||||| | |   |||||  | |||||||
Db   48073 gcctggctaagattttgtatttttattagagaaggggttttgccatgttggccaagctgg 48132

Qy    7095 tctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgc 7154
           |||| ||||  |  |||||  |  ||||||||||||  ||||||||||||||||||||
Db   48133 tctcgaactcctaacctcaggtgatccacctgcctcagcctcccaaagtgctgagattat 48192

Qy    7155 agatgtgagccaccgtgcctggccagatttttcttttattctt 7197
           ||   ||||||||  ||||   || | ||| || ||  || ||
Db   48193 aggcatgagccactatgccccacctggtttctccttatttatt 48235


RESULT 11
A35030
ID   A35030 standard; DNA; 141589 BP.
XX
AC   A35030;
XX
DT   28-JUL-2000 (first entry)
XX
DE   Human adenosine receptor related polynucleotide SEQ ID NO:2719.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17712.
XX
PR   03-AUG-1998;   98US-0095212.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstriction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers -
XX
PS   Disclosure; Page 972-1003; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
```

```
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. A32313 to A35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC   185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC   differ from the previously named sequences. SEQ ID NO:11 to 1680 (A32323
CC   to A33992) are specifically claimed ONs from the present invention.
CC   N.B. Sequences given in the disclosure of the present invention do not
CC   match up with their corresponding SEQ ID NO: sequences given in the
CC   sequence listing.
XX
SQ   Sequence 141589 BP; 42856 A; 28938 C; 26863 G; 42932 T; 0 other;


  Query Match             2.3%;  Score 205.4;  DB 21;  Length 141589;
  Best Local Similarity   74.9%;  Pred. No. 1.7e-27;
  Matches  257;  Conservative    0;  Mismatches   86;  Indels    0;  Gaps    0;

Qy    6855 ttctctgttgtttggttcagggattttatttttaattttatgagacagggtcttgctctg 6914
           || | ||||  ||  ||     |||| |||||  |||| ||||||||||||  |||||
Db   47893 ttgtttgtttgtttatttgtttctttttttttttttttttttgagacagggtctcactctg 47952

Qy    6915 tcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggc 6974
           |  ||||||||||||||||||||| ||||||   ||||||||||||||   ||||  ||
Db   47953 ttgcccaggctggagtgcagtggtatgatctcggctcactgcagcctccatctccctggt 48012

Qy    6975 tcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccgt 7034
           |||||  || ||| || |||||||| || | ||||||||||||||||||  |||||
Db   48013 tcaagtgattctcatgtctcagcctcccgagtagctgggactacaggtgcatgccactac 48072

Qy    7035 gcctggctaacttttcatttttttttgtagggacggggtctcgttttgttgccaaagctgg 7094
           ||||||||||  |||  | |||||  ||| || ||||| | |   ||||| | |||||||
Db   48073 gcctggctaagattttgtatttttattagagaaggggttttgccatgttggccaagctgg 48132

Qy    7095 tctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattgc 7154
           |||| ||||  |  ||||| |  |||||||||||  ||||||||||||||||||||||
Db   48133 tctcgaactcctaacctcaggtgatccacctgcctcagcctcccaaagtgctgagattat 48192

Qy    7155 agatgtgagccaccgtgcctggccagatttttcttttattctt 7197
           ||   ||||||||  ||||   || | ||| || ||  || ||
Db   48193 aggcatgagccactatgccccacctggtttctccttatttatt 48235


RESULT  12
A35191/c
ID   A35191 standard; DNA; 1601 BP.
XX
AC   A35191;
XX
DT   28-JUL-2000  (first entry)
XX
DE   Human adenosine receptor related polynucleotide 2nd SEQ ID NO:65.
XX
KW   Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW   phosphorothioate; impaired respiration; inflammation; allergy;
KW   allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW   antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW   lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
KW   respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW   pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW   cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS   Homo sapiens.
XX
PN   WO200009525-A2.
XX
PD   24-FEB-2000.
XX
PF   03-AUG-1999;   99WO-US17712.
XX
PR   03-AUG-1998;   98US-0095212.
XX
PA   (UYEC-) UNIV EAST CAROLINA.
XX
PI   Nyce JW;
XX
DR   WPI; 2000-205971/18.
XX
PT   New antisense oligonucleotides useful for treating e.g. pulmonary
PT   vasoconstriction, inflammation, allergies, asthma, hypertension,
PT   bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT   cancers -
XX
PS   Disclosure; Page 1232; 1343pp; English.
XX
CC   The present invention describes a new composition comprising an antisense
CC   oligonucleotide (ON) with low adenosine (up to 15%), which targets
CC   nucleic acids involved in bronchoconstriction, allergies, and/or
CC   inflammation. The ON can have antiinflammatory, antiallergic,
CC   antiasthmatic, cytostatic and analgesic activities. The compositions are
CC   useful for the treatment of diseases associated with inflammation,
CC   impaired airways, including lung disease and diseases whose secondary
CC   effects afflict the lungs of a subject. They can be used for treating
CC   e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,
CC   impeded respiration, respiratory distress syndrome, pain, cystic
CC   fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC   pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC   carcinomas, and cancers which may metastasise to the lungs, including
CC   breast and prostate cancer. The reduction of the adenosine content of the
CC   ONs reduces side effects. The A-containing ONs break down with the
CC   release of deoxyadenosine which activates adenosine receptors causing
CC   bronchoconstriction and inflammation. A32313 to A35312 represent the
CC   nucleotide sequences given in the sequence listing from the present
CC   invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC   185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC   differ from the previously named sequences. SEQ ID NO:11 to 1680 (A32323
CC   to A33992) are specifically claimed ONs from the present invention.
CC   N.B. Sequences given in the disclosure of the present invention do not
CC   match up with their corresponding SEQ ID NO: sequences given in the
CC   sequence listing.
XX
SQ   Sequence 1601 BP; 445 A; 368 C; 356 G; 432 T; 0 other;


  Query Match             2.3%;  Score 204.2;  DB 21;  Length 1601;
  Best Local Similarity   82.1%;  Pred. No. 1.2e-27;
  Matches  247;  Conservative    0;  Mismatches   53;  Indels    1;  Gaps    1;

Qy    6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
           ||||||||||| |||| |||||||| |||||||||||| |||||||||||| ||| ||
Db    1473 TTTTATTTTTATTTTTTTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGTACAGGGG 1414

Qy    6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
           ||||||   |||||||||||| | ||   |||||||| ||||||||| ||| |||||||||
Db    1413 CATGATATCAGCTCACTGCAACTTCCATCTCCTGGGTTCAAGCAATTCTCTTGCCTCAGC 1354

Qy    6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
           || |||| |||||||||| |||||||||||||||| |||| |||||| |||| | ||||
Db    1353 CTCCCAAGTAGCTGGGATTACAGGTGCGCGCCACCATGCCCGGCTAA-TTTTTGTATTTT 1295

Qy    7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
           | |||| |||||||| ||    ||||| | | ||||||    ||||  || |||||||
Db    1294 TAGTAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTTAGGAACTCCTGACCTCAAGTG 1235
```

```
Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| ||||  ||||||||||||||||| |||| |||  |||||||||| ||| |||
Db   1234 ATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGC 1175

Qy   7178 c 7178
          |
Db   1174 C 1174
```

RESULT 13
Z46814/c
ID Z46814 standard; cDNA; 1601 BP.
XX
AC Z46814;
XX
DT 31-MAR-2000 (first entry)
XX
DE Interleukin-10 (IL-10) cDNA sequence.
XX
KW Interleukin-10; IL-10; atopic dermatitis; allergic dermatitis;
KW SLE; EB viral infection; lymphoma; antisense; ss.
XX
OS Homo sapiens.
XX
PN WO9967388-A1.
XX
PD 29-DEC-1999.
XX
PF 22-JUN-1999; 99WO-JP03315.
XX
PR 24-JUN-1998; 98JP-0177188.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
XX
PI Arima H, Tsuchiya S, Hirata T, Akiyama K, Goto T;
XX
DR WPI; 2000-106294/09.
XX
PT Antisense oligonucleotide for inhibiting expression of IL-10 protein,
PT useful as active ingredient in remedies for atopic dermatitis -
XX
PS Disclosure; Page 19-20; 22pp; Japanese.
XX
CC The invention provides antisense oligonucleotides (Z46806-812) for
CC inhibiting expression of interleukin-10 (IL-10) protein. A composition
CC which is formulated with the antisense oligonucleotide or its derivatives
CC as active ingredient can be used for treating atopic dermatitis, allergic
CC dermatitis, SLE, EB viral infection or lymphoma. The present sequence
CC represents a IL-10 cDNA sequence.
XX
SQ Sequence 1601 BP; 445 A; 368 C; 356 G; 432 T; 0 other;

Query Match 2.3%; Score 204.2; DB 21; Length 1601;
Best Local Similarity 82.1%; Pred. No. 1.2e-27;
Matches 247; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

```
Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          |||||||||||  |||| |||||||| |||||||||||||  |||||||||||| ||| ||
Db   1473 TTTTATTTTTATTTTTTTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGTACAGGGG 1414

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
          ||||||   |||||||||||  | ||   |||||||| |||||||| ||| |||||||||
Db   1413 CATGATATCAGCTCACTGCAACTTCCATCTCCTGGGTTCAAGCAATTCTCTTGCCTCAGC 1354

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || |||| |||||||||| |||||||||||||||| |||| |||||| ||||  | ||||
Db   1353 CTCCCAAGTAGCTGGGATTACAGGTGCGCGCCACCATGCCCGGCTAA-TTTTTGTATTTT 1295

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| |||||||| ||    ||||| | | ||||||    ||||  || |||||||
Db   1294 TAGTAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTTAGGAACTCCTGACCTCAAGTG 1235
```

```
Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| ||||  ||||||||||||||||| |||| |||  |||||||||| ||| |||
Db   1234 ATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGC 1175

Qy   7178 c 7178
          |
Db   1174 C 1174
```

RESULT 14
Q10207/c
ID Q10207 standard; cDNA; 1618 BP.
XX
AC Q10207;
XX
DT 19-MAR-1991 (first entry)
XX
DE pH15C insert containing Human cytokine synthesis inhibitory
DE factor.
XX
KW Delayed-type hypersensitivity; DTH; leishmaniasis; parasite;
KW MHC-associated autoimmune disease; interferon-gamma; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 33..566
FT /*tag= a
XX
PN EP405980-A.
XX
PD 02-JAN-1991.
XX
PF 28-JUN-1990; 90EP-0307091.
XX
PR 20-DEC-1989; 89US-0453951.
PR 28-JUN-1989; 89US-0372667.
XX
PA (SCHE) SCHERING CORP.
XX
PI Mosmann TR, Moore KW, Bond MW, Vleira PJM;
XX
DR WPI; 1991-009290/02.
DR R-PSDB; R10158.
XX
PT Mammalian cytokine synthesis inhibitory factors - capable of
PT inhibiting synthesis of cytokine(s) associated with delayed-type
PT hypersensitivity and useful in treatment of e.g. leishmaniasis
XX
PS Disclosure; Fig 4; 31pp; English.
XX
CC The gene product may be used in treatment of diseases associated
CC with MHC-linked immune response, supressing a cell mediated or
CC humoral immune response. It may specifically be used to treat
CC delayed-type hyper-sensitivity, leishmaniasis, and immune disorders.
XX
SQ Sequence 1618 BP; 461 A; 369 C; 356 G; 432 T; 0 other;

Query Match 2.3%; Score 204.2; DB 12; Length 1618;
Best Local Similarity 82.1%; Pred. No. 1.2e-27;
Matches 247; Conservative 0; Mismatches 53; Indels 1; Gaps 1;

```
Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          |||||||||||  |||| |||||||| |||||||||||||  |||||||||||| ||| ||
Db   1473 TTTTATTTTTATTTTTTTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGTACAGGGG 1414

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
          ||||||   |||||||||||  | ||   |||||||| |||||||| ||| |||||||||
Db   1413 CATGATATCAGCTCACTGCAACTTCCATCTCCTGGGTTCAAGCAATTCTCTTGCCTCAGC 1354

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
```

```
          || |||| ||||||||| |||||||||||||||| |||| |||||| |||| | ||||
Db   1353 CTCCCAAGTAGCTGGGATTACAGGTGCGCGCCACCATGCCCGGCTAA-TTTTTGTATTTT 1295

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| |||||||| ||    ||||| | | ||||||    ||||  || |||||||
Db   1294 TAGTAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTTAGGAACTCCTGACCTCAAGTG 1235

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| |||| |||||||||||||||| |||| ||| |||||||||| ||| |||
Db   1234 ATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGC 1175

Qy   7178 c 7178
          |
Db   1174 C 1174
```

```
RESULT 15
Q46958/c
ID   Q46958 standard; cDNA; 1618 BP.
XX
AC   Q46958;
XX
DT   25-JAN-1994  (first entry)
XX
DE   Human cytokine synthesis inhibitory factor clone pH15C.
XX
KW   Mammalian cytokine synthesis inhibitory factor; CSIF; interleukin 10;
KW   IL-10; immune system imbalance; human T cell; Leischmaniasis;
KW   rheumatoid arthritis; systemic lupus erythematosus; thyroiditis;
KW   myasthenia gravis; insulin-dependent diabetes mellitus; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   CDS             31..567
FT                   /*tag= a
FT   mat_peptide     85..564
FT                   /*tag= b
FT                   /product= mature_CSIF
XX
PN   US5231012-A.
XX
PD   27-JUL-1993.
XX
PF   28-JUN-1989;   89US-0372667.
XX
PR   28-JUN-1989;   89US-0372667.
PR   20-DEC-1989;   89US-0453951.
PR   06-AUG-1990;   90US-0546235.
PR   20-JUL-1992;   92US-0917806.
XX
PA   (SCHE ) SCHERING CORP.
XX
PI   Bond MW,  Moore KW,  Mosmann TR,  Vieira PJM;
XX
DR   WPI; 1993-249726/31.
DR   P-PSDB; R39714.
XX
PT   Genes and proteins encoding cytokine synthesis inhibitory factors
PT   - useful in treating diseases associated with cytokine
PT   imbalances, e.g. parasitic infections and auto-immune disorders
XX
PS   Claim 1; Fig 4; 23pp; English.
XX
CC   A human T cell cDNA library was screened with probes based on the
CC   murine CSIF gene. Two clones carrying plasmids pH5C and pH15C were
CC   identified. (These two expression vectors are claimed). The CSIF
CC   polypeptide they encode inhibits synthesis of cytokines associated
CC   with delayed type hypersensitivity responses. CSIF (also called IL-
CC   10) can be used to treat diseases associated with cytokine
CC   imbalances, such as leichmaniasis and MHC-associated autoimmune
CC   diseases caused by excessive production of interferon-gamma, e.g.
CC   rheumatoid arthritis, SLE, IDDM, myasthenia gravis and thyroiditis.
XX
SQ   Sequence 1618 BP; 463 A; 367 C; 356 G; 432 T; 0 other;
```

```
  Query Match             2.3%;  Score 204.2;  DB 14;  Length 1618;
  Best Local Similarity   82.1%;  Pred. No. 1.2e-27;
  Matches  247;  Conservative    0;  Mismatches   53;  Indels    1;  Gaps    1;

Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          |||||||||||| |||| |||||||| |||||||||||||| |||||||||||||| ||| ||
Db   1473 TTTTATTTTTATTTTTTTGAGACAGAGTCTTGCTCTGTCACCCAGGCTGGAGTACAGGGG 1414

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
          ||||||   |||||||||||| | ||   |||||||| ||||||||| ||| |||||||||
Db   1413 CATGATATCAGCTCACTGCAACTTCCATCTCCTGGGTTCAAGCAATTCTCTTGCCTCAGC 1354

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || |||| ||||||||| |||||||||||||||||| |||| |||||| ||||  | ||||
Db   1353 CTCCCAAGTAGCTGGGATTACAGGTGCGCGCCACCATGCCCGGCTAA-TTTTTGTATTTT 1295

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| |||||||| ||    ||||| | | ||||||    ||||  || |||||||
Db   1294 TAGTAGAGACGGGGTTTCACCATGTTGACCAGGCTGGTTAGGAACTCCTGACCTCAAGTG 1235

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| ||||  |||||||||||||||| |||| |||  |||||||||| ||| |||
Db   1234 ATCCACCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCCGGC 1175

Qy   7178 c 7178
          |
Db   1174 C 1174
```

```
Search completed: April  6, 2001, 00:50:49
Job time: 17904 sec
```

GenCore version 4.5

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:35:45 ; Search time 152.82 Seconds
(without alignments)
9541.804 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 280836 seqs, 80580151 residues

Total number of hits satisfying chosen parameters: 561672

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6_COMB.seq:*
4: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq:*
5: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	210.6	2.3	1811	1	US-08-848-252-1	Sequence 1, Appli
c	2	209.8	2.3	35060	3	US-08-814-095-7	Sequence 7, Appli
c	3	201.6	2.2	53526	3	US-08-658-136-2	Sequence 2, Appli
c	4	201.6	2.2	53577	3	US-08-658-136-1	Sequence 1, Appli
	5	199	2.2	1988	2	US-08-257-963B-11	Sequence 11, Appl
	6	199	2.2	1988	4	PCT-US95-07201-11	Sequence 11, Appl
	7	199	2.2	22481	4	PCT-US95-07201-43	Sequence 43, Appl
c	8	195.8	2.2	1600	2	US-08-487-113D-117	Sequence 117, App
c	9	195.8	2.2	1600	2	US-08-720-420A-117	Sequence 117, App
c	10	195.8	2.2	7210	2	US-08-257-963B-10	Sequence 10, Appl
c	11	195.8	2.2	7210	4	PCT-US95-07201-10	Sequence 10, Appl
c	12	195.8	2.2	22481	4	PCT-US95-07201-43	Sequence 43, Appl
c	13	195.2	2.2	7676	1	US-08-451-777A-7	Sequence 7, Appli
c	14	195.2	2.2	7676	2	US-08-451-778A-7	Sequence 7, Appli
c	15	195.2	2.2	7676	2	US-08-998-208-7	Sequence 7, Appli
c	16	195.2	2.2	7676	4	PCT-US95-06743-7	Sequence 7, Appli
	17	194	2.1	20303	1	US-08-370-975B-6	Sequence 6, Appli
	18	194	2.1	26764	1	US-08-370-975B-1	Sequence 1, Appli
	19	194	2.1	87350	3	US-08-781-891-79	Sequence 79, Appl
	20	193.8	2.1	282	1	US-08-133-629-8	Sequence 8, Appli
c	21	193.6	2.1	4823	2	US-08-457-254-5	Sequence 5, Appli
c	22	193.6	2.1	4823	2	US-08-484-257-20	Sequence 20, Appl
c	23	193.6	2.1	4823	3	US-08-999-927-5	Sequence 5, Appli
c	24	193.6	2.1	4823	4	PCT-US94-08806-28	Sequence 28, Appl
c	25	193.6	2.1	4823	4	PCT-US95-01829-5	Sequence 5, Appli
c	26	193.6	2.1	4823	4	PCT-US95-16626-5	Sequence 5, Appli
c	27	193.2	2.1	4220	1	US-08-832-883-66	Sequence 66, Appl
c	28	193.2	2.1	4220	2	US-08-832-877-66	Sequence 66, Appl
	29	193.2	2.1	53526	3	US-08-658-136-2	Sequence 2, Appli
	30	193.2	2.1	53577	3	US-08-658-136-1	Sequence 1, Appli
	31	193.2	2.1	56516	2	US-08-996-306-1	Sequence 1, Appli
c	32	193	2.1	1988	2	US-08-257-963B-11	Sequence 11, Appl
c	33	193	2.1	1988	4	PCT-US95-07201-11	Sequence 11, Appl
	34	192	2.1	11531	1	US-08-068-945A-1	Sequence 1, Appli
	35	192	2.1	11531	1	US-08-442-806-1	Sequence 1, Appli
c	36	191.8	2.1	17327	1	US-07-906-871-15	Sequence 15, Appl
c	37	191.6	2.1	176373	3	US-09-128-155-17	Sequence 17, Appl
c	38	191.4	2.1	87350	3	US-08-781-891-79	Sequence 79, Appl
	39	191.2	2.1	13865	3	US-09-009-217-11	Sequence 11, Appl
	40	191.2	2.1	13865	3	US-09-009-656-11	Sequence 11, Appl
c	41	190.4	2.1	246240	2	US-08-724-394A-20	Sequence 20, Appl
c	42	190.4	2.1	246240	2	US-08-724-394A-21	Sequence 21, Appl
c	43	190.4	2.1	246240	2	US-08-724-394A-22	Sequence 22, Appl
	44	189.8	2.1	28994	3	US-08-884-324-14	Sequence 14, Appl
	45	188.8	2.1	3607	2	US-08-629-001A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1
US-08-848-252-1/c
; Sequence 1, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,052
; FILING DATE:
; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McDiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1811 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; LIBRARY: Human
; CLONE: Signal Transductor CD24
; FEATURE:
; NAME/KEY: sig_peptide

```
;      LOCATION:  57..134
;    FEATURE:
;      NAME/KEY:  mat_peptide
;      LOCATION:  135..296
;    FEATURE:
;      NAME/KEY:  CDS
;      LOCATION:  57..299
US-08-848-252-1

  Query Match             2.3%;  Score 210.6;  DB 1;  Length 1811;
  Best Local Similarity   73.9%;  Pred. No. 6.2e-36;
  Matches  283;  Conservative    0;  Mismatches   94;  Indels    6;  Gaps    1;

Qy    4275 aattattagtgaaagtatgtttataataatatacttctgtgtggctaggcgtggtggctc 4334
           ||  | || | |||          |||||  |  | | |   ||||| |||| || |||||
Db    1469 AAAAAGTATTCAAACATATGCAGAAATAAAAAGCATTTTGATGGCTGGGCGCGGGGGCTC 1410

Qy    4335 acgcctgcaatcccagcactttgggaggcagaggcaggcagagcacttgaggtcaggaat 4394
           | ||||| |||||||||||||||||||| |||||  ||  || ||| ||||||||||| |
Db    1409 AAGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCGGGTGGATCACCTGAGGTCAGGAGT 1350

Qy    4395 tcgagatcagcctggccaacatgaggaaaccccgtctctactaaaaatacaaaaatcagc 4454
           |||| | |||||||| |||||||  ||||||| |||||||||||||| |||||||| |||
Db    1349 TCGAAACCAGCCTGGGCAACATGGTGAAACCCTGTCTCTACTAAAATTACAAAAATTAGC 1290

Qy    4455 caggcatggtggcaggtgtctgtaatcccagctacttgggaggctgaggcaggagaatc- 4513
           | ||||||||||||||| ||||||||||||||||||||||||||||||||||||||||
Db    1289 CGGGCATGGTGGCAGGTGCCTGTAATCCCAGCTACTTGGGAGGCTGAGGCAGGAGAATCG 1230

Qy    4514 -----agaggggaggcggaggttgcagtgagccaagatcacgccactacaccccagccta 4568
                  |||||||  | | | || ||| |  ||||| | ||||| ||| || ||||
Db    1229 CTTGAGCCCGGGAGGCAAATGGTACACTGATCTGAGATCGCACCACTGCACTCCGGCCTG 1170

Qy    4569 ggtgacaaagcgagacttctcaaatattaacaataataatatactatgtgtcattataca 4628
           || ||||||| ||||||    ||| | ||| |||||||| |  |  | ||  ||||  |
Db    1169 GGCGACAAAGTGAGACTGTCTAAAAAATAATAATAATAAAAAGCATTTTGAAATTAGTCG 1110

Qy    4629 tgatgattattattttatcattt 4651
           || | | |   ||| ||   |||
Db    1109 TGGTCAATGCAATTCTACTCTTT 1087


RESULT   2
US-08-814-095-7/c
; Sequence 7, Application US/08814095
; Patent No. 6025183
;  GENERAL INFORMATION:
;    APPLICANT:  Soreq, Hermona
;    APPLICANT:  Zakut, Haim
;    APPLICANT:  Shani, Moshe
;    TITLE OF INVENTION:  TRANSGENIC ANIMAL ASSAY SYSTEM FOR
;    TITLE OF INVENTION:  ANTI-CHOLINESTERASE SUBSTNACES
;    NUMBER OF SEQUENCES:  7
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  KOHN & ASSOCIATES
;      STREET:  30500 No. 6025183thwestern Highway, Suite 410
;      CITY:  Farmington Hills
;      STATE:  Michigan
;      COUNTRY:  U.S.
;      ZIP:  48334
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/814,095
;      FILING DATE:
;      CLASSIFICATION:  800
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Montgomery, Ilene N.
;      REGISTRATION NUMBER:  38,972
;      REFERENCE/DOCKET NUMBER:  2391.00066
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (248) 539-5050
;      TELEFAX:  (248) 539-5055
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  35060 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  other nucleic acid
;      DESCRIPTION:  /desc = "Cosmid including ACHE
;      DESCRIPTION:  promotor, ACHE gene and ARS gene"
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  NO
;    ORIGINAL SOURCE:
;      ORGANISM:  Homo sapiens
;    POSITION IN GENOME:
;      CHROMOSOME/SEGMENT:  7q22
;    FEATURE:
;      NAME/KEY:  promoter
;      LOCATION:  4089..22464
;      OTHER INFORMATION:  /function= "ACHE Promotor"
;      OTHER INFORMATION:  /standard_name= "ACHE Promotor"
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  22465..22537
;      OTHER INFORMATION:  /function= "non-translated"
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 1
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  24090..25177
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /function= "(translation start:
;      OTHER INFORMATION:  24110)"
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 2
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  25524..26009
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 3
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  27005..27274
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 4
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  27255..28007
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 5
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  27385..27387
;    FEATURE:
;      NAME/KEY:  exon
;      LOCATION:  28008..28129
;      IDENTIFICATION METHOD:  experimental
;      OTHER INFORMATION:  /evidence= EXPERIMENTAL
;      OTHER INFORMATION:  /gene= "ACHE"
;      OTHER INFORMATION:  /number= 6
;    FEATURE:
;      NAME/KEY:  terminator
;      LOCATION:  28129..28131
```

```
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (34528..34895)
;       OTHER INFORMATION:  /function= "arsenite resistance
;       OTHER INFORMATION:  gene"
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 1
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (34092..34358)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 2
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33779..33963)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 3
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33493..33591)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 4
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (33297..33408)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 5
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32959..33094)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 6
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32569..32628)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 7
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (32386..32468)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 8
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31894..32080)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 9
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31363..31534)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 10
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (31131..31284)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 11
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30816..31011)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 12
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30470..30626)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 13
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (30187..30274)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 14
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (29945..30073)
;       OTHER INFORMATION:  /gene= "AR"
;       OTHER INFORMATION:  /number= 15
;     FEATURE:
;       NAME/KEY:  exon
;       LOCATION:  complement (29664..29856)
;       OTHER INFORMATION:  /gene= "ARS"
;       OTHER INFORMATION:  /number= 16
US-08-814-095-7


  Query Match             2.3%;  Score 209.8;  DB 3;  Length 35060;
  Best Local Similarity   81.5%;  Pred. No. 3.6e-35;
  Matches  255;  Conservative    0;  Mismatches   57;  Indels    1;  Gaps    1;

Qy   6879 tttatttttaattttatgagacagggtcttgctctgtccccaggctggagtgcagtggc 6938
          | | |||||  ||||| | ||||||||||||||||||| ||||||||| ||| |||||||
Db   8387 TATTTTTTTCTTTTTAAGGGACAGGGTCTTGCTCTGTCACCCAGGCTGTAGTACAGTGGC 8328

Qy   6939 atgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagcc 6998
          || |||||||||||||||||||| |||||||||||||||  | |||||  ||||||||
Db   8327 ATAATCATAGCTCACTGCAGCCTCAAACTCCTGGGCTCAAGTGAGCCTCCCACCTCAGCC 8268

Qy   6999 ttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7058
          | || | |||||||||||||||| | ||||||  ||  |||||| ||||    |||||
Db   8267 TCCCCAGTAGCTGGGACTACAGGTGTGTGCCACCCCGCTGGGCTAA-TTTTTGCATTTTT 8209

Qy   7059 tgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagcaa 7118
          ||||| || ||||||| | | ||||||| | ||||||||| |||| |||  |||||||||
Db   8208 TGTAGAGAAGGGGTCTTGCTATGTTGCCCAGGCTGGTCTCGAACTCGTGAGCTCAAGCAA 8149

Qy   7119 tccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggcc 7178
          ||  || ||||  |||||||||||||||| |||| |||| ||||||||| ||| |||| |
Db   8148 TCTGCCCGCCTCAGCCTCCCAAAGTGCTGGGATTACAGAAGTGAGCCACTGTGTCTGGTC 8089

Qy   7179 agatttttctttt 7191
               ||||  |||
Db   8088 TACGTTTTTATTT 8076


RESULT   3
US-08-658-136-2/c
; Sequence 2, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;    APPLICANT:  KLINGER, KATHERINE W
;    APPLICANT:  LANDES, GREGORY M
;    APPLICANT:  BURN, TIMOTHY C
;    APPLICANT:  CONNORS, TIMOTHY D
;    APPLICANT:  DACKOWSKI, WILLIAM
;    APPLICANT:  GERMINO, GREGORY
;    APPLICANT:  QIAN, FENG
;    TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GENZYME CORPORATION
;      STREET:  ONE MOUNTAIN ROAD
;      CITY:  FRAMINGHAM
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,136
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LASSEN, ELIZABETH
```

```
;       REGISTRATION NUMBER:  31,845
;       REFERENCE/DOCKET NUMBER:  GEN4-17.8
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  508-872-8400
;       TELEFAX:  508-872-5415
;   INFORMATION FOR SEQ ID NO:  2:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  53526 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
;     MOLECULE TYPE:  DNA (genomic)
US-08-658-136-2


  Query Match             2.2%;  Score 201.6;  DB 3;  Length 53526;
  Best Local Similarity   73.3%;  Pred. No. 2.5e-33;
  Matches  258;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy   6845 aaggtggagtttctctgttgtttggttcagggatttttatttttaatttttatgagacaggg 6904
          ||||    |   ||  | | ||   |   ||  |||| |||||  ||||  ||||  | |
Db  42800 AAGGCTCTGAGCCTGAGCTTTTCTTTGTGGGATTTTTTTTTTTTTTTTTGGAGATGGAG 42741

Qy   6905 tcttgctctgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcat 6964
          |||||||| ||  |||||| ||||||||||||||   |||   ||||||||||  |||
Db  42740 TCTTGCTCCGTTGCCCAGGATGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAAGCTCCG 42681

Qy   6965 actcctgggctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgc 7024
          ||||||||  ||| || || ||||||||||||||  | | ||||||||||||||| ||
Db  42680 CCTCCTGGGTTCATGCCATTCTCCTGCCTCAGCCTCTCGAGTAGCTGGGACTACAGGCGC 42621

Qy   7025 gcgccaccgtgcctggctaacttttcatttttttttgtagggacggggtctcgttttgttg 7084
          |||||||  ||||  |||||  ||||  | |||  | ||||  | ||||||  || || |||||
Db  42620 CCGCCACCATGCCCAGCTAAGTTTTTGTATTTGTAGTAGAGTCGGGGTTTCATTGTGTTG 42561

Qy   7085 ccaaagctggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtg 7144
          | | |||||||||  ||||  ||  |||||||  |||  ||  |||| ||||||||||||||
Db  42560 GCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCTGCCCGCCTCGGCCTCCCAAAGTG 42501

Qy   7145 ctgagattgcagatgtgagccaccgtgcctggccagatttttcttttattct 7196
          |||  ||||  |||  ||||||||||  ||||||||   |||    |||  | |
Db  42500 CTGGGATTACAGGCGTGAGCCACCGCGCCTGGCCCTTTTTAATGTTTTATAT 42449


RESULT   4
US-08-658-136-1/c
; Sequence 1, Application US/08658136
; Patent No. 6071717
;  GENERAL INFORMATION:
;    APPLICANT:  KLINGER, KATHERINE W
;    APPLICANT:  LANDES, GREGORY M
;    APPLICANT:  BURN, TIMOTHY C
;    APPLICANT:  CONNORS, TIMOTHY D
;    APPLICANT:  DACKOWSKI, WILLIAM
;    APPLICANT:  GERMINO, GREGORY
;    APPLICANT:  QIAN, FENG
;    TITLE OF INVENTION:  POLYCYSTIC KIDNEY DISEASE GENE
;    NUMBER OF SEQUENCES:  58
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  GENZYME CORPORATION
;      STREET:  ONE MOUNTAIN ROAD
;      CITY:  FRAMINGHAM
;      STATE:  MASSACHUSETTS
;      COUNTRY:  USA
;      ZIP:  01701
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/658,136
;    FILING DATE:
;       CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  LASSEN, ELIZABETH
;      REGISTRATION NUMBER:  31,845
;      REFERENCE/DOCKET NUMBER:  GEN4-17.8
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  508-872-8400
;      TELEFAX:  508-872-5415
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  53577 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-658-136-1


  Query Match             2.2%;  Score 201.6;  DB 3;  Length 53577;
  Best Local Similarity   73.3%;  Pred. No. 2.5e-33;
  Matches  258;  Conservative    0;  Mismatches   94;  Indels    0;  Gaps    0;

Qy   6845 aaggtggagtttctctgttgtttggttcagggatttttatttttaatttttatgagacaggg 6904
          ||||    |   ||  | | ||   |   ||  |||| |||||  ||||  ||||  | |
Db  42795 AAGGCTCTGAGCCTGAGCTTTTCTTTGTGGGATTTTTTTTTTTTTTTTTGGAGATGGAG 42736

Qy   6905 tcttgctctgtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcat 6964
          |||||||| ||  |||||| ||||||||||||||   |||   ||||||||||  |||
Db  42735 TCTTGCTCCGTTGCCCAGGATGGAGTGCAGTGGCGCAATCTCGGCTCACTGCAAGCTCCG 42676

Qy   6965 actcctgggctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgc 7024
          ||||||||  ||| || || ||||||||||||||  | | ||||||||||||||| ||
Db  42675 CCTCCTGGGTTCATGCCATTCTCCTGCCTCAGCCTCTCGAGTAGCTGGGACTACAGGCGC 42616

Qy   7025 gcgccaccgtgcctggctaacttttcatttttttttgtagggacggggtctcgttttgttg 7084
          |||||||  ||||  |||||  ||||  | |||  | ||||  | ||||||  || || |||||
Db  42615 CCGCCACCATGCCCAGCTAAGTTTTTGTATTTGTAGTAGAGTCGGGGTTTCATTGTGTTG 42556

Qy   7085 ccaaagctggtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtg 7144
          | | |||||||||  ||||  ||  |||||||  |||  ||  |||| ||||||||||||||
Db  42555 GCCAGGCTGGTCTCGAACTCCTGACCTCAAGTGATCTGCCCGCCTCGGCCTCCCAAAGTG 42496

Qy   7145 ctgagattgcagatgtgagccaccgtgcctggccagatttttcttttattct 7196
          |||  ||||  |||  ||||||||||  ||||||||   |||    |||  | |
Db  42495 CTGGGATTACAGGCGTGAGCCACCGCGCCTGGCCCTTTTTAATGTTTTATAT 42444


RESULT   5
US-08-257-963B-11
; Sequence 11, Application US/08257963B
; Patent No. 5840686
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, S.
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
;    TITLE OF INVENTION:  BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
;    TITLE OF INVENTION:  AND EXPRESSING THE PROTEIN
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
```

```
;      APPLICATION NUMBER:  US/08/257,963B
;      FILING DATE:
;      CLASSIFICATION:  514
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/952,796
;      FILING DATE:  24-SEPT-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36434
;      REFERENCE/DOCKET NUMBER:  20264126US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1988 Base Pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;    FEATURE:
;      NAME/KEY:  JT108
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  2 kb PCR product using
;      OTHER INFORMATION:  primers, SEQ ID: 13 and 14.
US-08-257-963B-11


  Query Match             2.2%;  Score 199;  DB 2;  Length 1988;
  Best Local Similarity   78.8%;  Pred. No. 1.9e-33;
  Matches  246;  Conservative    0;  Mismatches   65;  Indels    1;  Gaps    1;

Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          | || ||||    ||| |||||| | |||| ||| |||| || ||||||||||||||||||
Db    541 TNTTTTTTTNNNCTTTCTGAGACGGAGTCTCGCTTTGTCNCCAAGGCTGGAGTGCAGTGG 600

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
            |||||  |||||||||||| ||||   ||||  || ||||| ||| ||||||||||||
Db    601 TGTGATCTCAGCTCACTGCAACCTCTGCCTCCCTGGTTCAAGGAATTCTCCTGCCTCAGC 660

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || |  | ||||||||| ||||||  |  |||| | ||||  ||||| |||  | ||||
Db    661 CTCCAGAGTAGCTGGGATTACAGGCACCTGCCATCATGCCCAGCTAA-TTTTTGTATTTT 719

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| ||||||||  ||    |||||||  | ||||||||||||||   || |||||||
Db    720 TAGTAGAGACGGGGTTTCACCATGTTGCCTAGGCTGGTCTCAAACTCCCGGGCTCAAGCG 779

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| |||||||||||||||||||| |||| |||  |||||||||| |||||||
Db    780 ATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCTGGC 839

Qy   7178 cagatttttctt 7189
          |||| |  ||||
Db    840 CAGAATAATCTT 851


RESULT   6
PCT-US95-07201-11
; Sequence 11, Application PC/TUS9507201
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, Sofia
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION GENOMIC
;    TITLE OF INVENTION:  ORGANIZATION AND SEQUENCE OF THE PEDF GENE
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07201
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/367,841
;      FILING DATE:  30-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/257,963
;      FILING DATE:  07-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/952,796
;      FILING DATE:  24-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36434
;      REFERENCE/DOCKET NUMBER:  20264126PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  11:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1988 Base Pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;    FEATURE:
;      NAME/KEY:  JT8A
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  2 kb PCR product using
;      OTHER INFORMATION:  primers, SEQ ID: 13 and 14; Also referred
;      OTHER INFORMATION:  to as JT108
PCT-US95-07201-11


  Query Match             2.2%;  Score 199;  DB 4;  Length 1988;
  Best Local Similarity   78.8%;  Pred. No. 1.9e-33;
  Matches  246;  Conservative    0;  Mismatches   65;  Indels    1;  Gaps    1;

Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          | || ||||    ||| |||||| | |||| ||| |||| || ||||||||||||||||||
Db    541 TNTTTTTTTNNNCTTTCTGAGACGGAGTCTCGCTTTGTCNCCAAGGCTGGAGTGCAGTGG 600

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
            |||||  |||||||||||| ||||   ||||  || ||||| ||| ||||||||||||
Db    601 TGTGATCTCAGCTCACTGCAACCTCTGCCTCCCTGGTTCAAGGAATTCTCCTGCCTCAGC 660

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || |  | ||||||||| ||||||  |  |||| | ||||  ||||| |||  | ||||
Db    661 CTCCAGAGTAGCTGGGATTACAGGCACCTGCCATCATGCCCAGCTAA-TTTTTGTATTTT 719

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| ||||||||  ||    |||||||  | ||||||||||||||   || |||||||
Db    720 TAGTAGAGACGGGGTTTCACCATGTTGCCTAGGCTGGTCTCAAACTCCCGGGCTCAAGCG 779

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| |||||||||||||||||||| |||| |||  |||||||||| |||||||
Db    780 ATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCTGGC 839
```

```
Qy   7178 cagatttttctt 7189
          |||| |  ||||
Db    840 CAGAATAATCTT 851

.
RESULT   7
PCT-US95-07201-43
; Sequence 43, Application PC/TUS9507201
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, Sofia
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION GENOMIC
;    TITLE OF INVENTION:  ORGANIZATION AND SEQUENCE OF THE PEDF GENE
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07201
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/367,841
;      FILING DATE:  30-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/257,963
;      FILING DATE:  07-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/952,796
;      FILING DATE:  24-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36434
;      REFERENCE/DOCKET NUMBER:  20264126PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22481 Base Pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Genomic DNA
;    FEATURE:
;      NAME/KEY:  Pl-147
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  full length genomic
;      OTHER INFORMATION:  sequence for PEDF plus flanking sequences.
PCT-US95-07201-43


  Query Match             2.2%;  Score 199;  DB 4;  Length 22481;
  Best Local Similarity   78.8%;  Pred. No. 6e-33;
  Matches  246;  Conservative    0;  Mismatches   65;  Indels    1;  Gaps    1;

Qy   6878 ttttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagtgg 6937
          | || ||||    ||| |||||| | |||| ||| |||| || ||||||||||||||||||
Db  15138 TNTTTTTTTNNNCTTTCTGAGACGGAGTCTCGCTTTGTCNCCAAGGCTGGAGTGCAGTGG 15197

Qy   6938 catgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctcagc 6997
            |||||  |||||||||||| ||||   ||||  || ||||| ||| ||||||||||||
Db  15198 TGTGATCTCAGCTCACTGCAACCTCTGCCTCCCTGGTTCAAGGAATTCTCCTGCCTCAGC 15257

Qy   6998 cttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttttt 7057
          || |  | ||||||||| ||||||  |  |||| | ||||  |||||  ||||  | ||||
Db  15258 CTCCAGAGTAGCTGGGATTACAGGCACCTGCCATCATGCCCAGCTAA-TTTTTGTATTTT 15316

Qy   7058 ttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaagca 7117
          | |||| |||||||| ||     ||||||| | ||||||||||||||   || |||||||
Db  15317 TAGTAGAGACGGGGTTTCACCATGTTGCCTAGGCTGGTCTCAAACTCCCGGGCTCAAGCG 15376

Qy   7118 atccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctggc 7177
          ||||||| |||||||||||||||||||||| |||| |||  |||||||||| |||||||
Db  15377 ATCCACCCGCCTTGGCCTCCCAAAGTGCTGGGATTACAGGCGTGAGCCACCGCGCCTGGC 15436

Qy   7178 cagatttttctt 7189
          |||| |  ||||
Db  15437 CAGAATAATCTT 15448


RESULT   8
US-08-487-113D-117/c
; Sequence 117, Application US/08487113D
; Patent No. 5837822
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  ICAM-Related Materials and Methods
;    NUMBER OF SEQUENCES:  120
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,113D
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,754
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,852
;      FILING DATE:  05-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/009,266
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
;      FILING DATE:  05-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  No. 5837822and, Greta E.
;      REGISTRATION NUMBER:  35,302
;      REFERENCE/DOCKET NUMBER:  32744
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  117:
;    SEQUENCE CHARACTERISTICS:
```

```
;      LENGTH:  1600 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA
US-08-487-113D-117


  Query Match             2.2%;  Score 195.8;  DB 2;  Length 1600;
  Best Local Similarity   77.9%;  Pred. No. 8.5e-33;
  Matches  236;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy   6876 gattttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagt 6935
          | ||||  |||   | || ||||||||| ||||  ||||||| |||||||||||||||||
Db    784 GGTTTTTGTTTGTTTGTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 725

Qy   6936 ggcatgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctca 6995
          |||  ||||   |||||||||||||||   |||| ||| |||| | ||  ||||||||||
Db    724 GGCGGGATCTCGGCTCACTGCAGCCTCTGCCTCCCGGGTTCAAACGATTTTCCTGCCTCA 665

Qy   6996 gccttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttt 7055
          |||| || | ||||||||| ||||||   | ||||||  ||| |||||| |||    ||
Db    664 GCCTCCCGAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTACTT 605

Qy   7056 ttttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaag 7115
          ||| |||| |||||||| ||    ||||| | ||||||||||||||||| || |||||||
Db    604 TTTGGTAGAGACGGGGTTTCACCATGTTGGCCAAGCTGGTCTCAAACTTCTGACCTCAAG 545

Qy   7116 caatccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctg 7175
            | |  || |||||||||||||||||| || |||| |||  ||||||||| | ||| |
Db    544 TGAGCTGCCCGCCTTGGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCCACTGCGCCCG 485

Qy   7176 gcc 7178
          |||
Db    484 GCC 482


RESULT  9
US-08-720-420A-117/c
; Sequence 117, Application US/08720420A
; Patent No. 5989843
;  GENERAL INFORMATION:
;    APPLICANT:  Gallatin, W. Michael
;    APPLICANT:  Vazeux, Rosemay
;    TITLE OF INVENTION:  ICAM-Related Materials and Methods
;    NUMBER OF SEQUENCES:  120
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Marshall, O'Toole, Gerstein, Murray & Borun
;      STREET:  6300 Sears Tower, 233 South Wacker Drive
;      CITY:  Chicago
;      STATE:  Illinois
;      COUNTRY:  United States of America
;      ZIP:  60606-6402
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/720,420A
;      FILING DATE:
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/487,113
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/286,754
;      FILING DATE:  05-AUG-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/102,852
;      FILING DATE:  05-AUG-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/009,266
;      FILING DATE:  22-JAN-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/894,061
;      FILING DATE:  05-JUN-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/889,724
;      FILING DATE:  26-MAY-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/827,689
;      FILING DATE:  27-JAN-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Williams, Joseph A., Jr.
;      REGISTRATION NUMBER:  38,659
;      REFERENCE/DOCKET NUMBER:  33282
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (312) 474-6300
;      TELEFAX:  (312) 474-0448
;      TELEX:  25-3856
;  INFORMATION FOR SEQ ID NO:  117:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  1600 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
US-08-720-420A-117


  Query Match             2.2%;  Score 195.8;  DB 2;  Length 1600;
  Best Local Similarity   77.9%;  Pred. No. 8.5e-33;
  Matches  236;  Conservative    0;  Mismatches   67;  Indels    0;  Gaps    0;

Qy   6876 gattttatttttaattttatgagacagggtcttgctctgtcccccaggctggagtgcagt 6935
          | ||||  |||   | || ||||||||| ||||  ||||||| |||||||||||||||||
Db    784 GGTTTTTGTTTGTTTGTTTTGAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGT 725

Qy   6936 ggcatgatcatagctcactgcagcctcatactcctgggctcaagcaatcctcctgcctca 6995
          |||  ||||   |||||||||||||||   |||| ||| |||| | ||  ||||||||||
Db    724 GGCGGGATCTCGGCTCACTGCAGCCTCTGCCTCCCGGGTTCAAACGATTTTCCTGCCTCA 665

Qy   6996 gccttccaactagctgggactacaggtgcgcgccaccgtgcctggctaacttttcatttt 7055
          |||| || | ||||||||| ||||||   | ||||||  ||| |||||| |||    ||
Db    664 GCCTCCCGAGTAGCTGGGATTACAGGCATGTGCCACCACGCCCGGCTAATTTTTGTACTT 605

Qy   7056 ttttgtagggacggggtctcgttttgttgccaaagctggtctcaaacttgtggcctcaag 7115
          ||| |||| |||||||| ||    ||||| | ||||||||||||||||| || |||||||
Db    604 TTTGGTAGAGACGGGGTTTCACCATGTTGGCCAAGCTGGTCTCAAACTTCTGACCTCAAG 545

Qy   7116 caatccacctgccttggcctcccaaagtgctgagattgcagatgtgagccaccgtgcctg 7175
            | |  || |||||||||||||||||| || |||| |||  ||||||||| | ||| |
Db    544 TGAGCTGCCCGCCTTGGCCTCCCAAAGTGTTGGGATTACAGGCGTGAGCCACTGCGCCCG 485

Qy   7176 gcc 7178
          |||
Db    484 GCC 482


RESULT  10
US-08-257-963B-10/c
; Sequence 10, Application US/08257963B
; Patent No. 5840686
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, S.
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
;    TITLE OF INVENTION:  BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
;    TITLE OF INVENTION:  AND EXPRESSING THE PROTEIN
;    NUMBER OF SEQUENCES:  42
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan
;      STREET:  345 Park Avenue
```

```
;       CITY:  New York
;       STATE:  New York
;       COUNTRY:  USA
;       ZIP:  10154
;     COMPUTER READABLE FORM:
;       MEDIUM TYPE:  Floppy Disk
;       COMPUTER:  IBM PC Compatible
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  WORDPERFECT 5.1
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/08/257,963B
;       FILING DATE:
;       CLASSIFICATION:  514
;     PRIOR APPLICATION DATA:
;       APPLICATION NUMBER:  07/952,796
;       FILING DATE:  24-SEPT-1992
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  DOROTHY R. AUTH
;       REGISTRATION NUMBER:  36434
;       REFERENCE/DOCKET NUMBER:  20264126US1
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (212) 758-4800
;       TELEFAX:  (212) 751-6849
;   INFORMATION FOR SEQ ID NO:  10:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  7210 Base Pairs
;       TYPE:  Nucleic Acid
;       STRANDEDNESS:  Double
;       TOPOLOGY:  Unknown
;     MOLECULE TYPE:  Genomic DNA
;     ORIGINAL SOURCE:
;       ORGANISM:  Human
;     IMMEDIATE SOURCE:
;       LIBRARY:  DASH II
;     FEATURE:
;       NAME/KEY:  JT106
;       LOCATION:
;       IDENTIFICATION METHOD:
;       OTHER INFORMATION:  7.2 kb No. 5840686 1 fragments
;       OTHER INFORMATION:  Derived from human placental genomic DNA
US-08-257-963B-10


  Query Match             2.2%;  Score 195.8;  DB 2;  Length 7210;
  Best Local Similarity   73.7%;  Pred. No. 1.7e-32;
  Matches  263;  Conservative    0;  Mismatches   92;  Indels    2;  Gaps    1;

Qy   6854 tttctctgttgtttggttcagggattttatttttaatttttatgagacagggtcttgctct 6913
          ||||| | |    |  ||       ||| | ||   |||| |||| ||| ||||||||||
Db   1893 TTTCTGTATCTGGTCCTTTTTTTTTATTTTTATTCTTTTTTTTGAGGCAGAGTCTTGCTCT 1834

Qy   6914 gtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctggg 6973
          ||| |||||||||||||||||||| |||||    | ||||||| ||||   ||||||||
Db   1833 GTCGCCCAGGCTGGAGTGCAGTGGCGTGATCTCGGTTCACTGCAACCTCTGCCTCCTGGG 1774

Qy   6974 ctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccg 7033
           |||||||||  ||||||||||||||| |  | |||||| |||||||| || |||||| |
Db   1773 GTCAAGCAATTATCCTGCCTCAGCCTCCTGAGTAGCTGAGACTACAGGCGCCCGCCACTG 1714

Qy   7034 tgcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctg 7093
            ||  ||||| ||||   |||||  |||| |||||||| ||    | ||| | | ||||
Db   1713 CACCCAGCTAATTTTTGTATTTTTCAGTAGAGACGGGGTTTCACCATTTTGGCCAGGCTG 1654

Qy   7094 gtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattg 7153
          ||||  ||||  || ||||| |  |||||||  ||| ||||||||||||||| ||||
Db   1653 GTCTTGAACTCCTGACCTCATG--ATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTA 1596

Qy   7154 cagatgtgagccaccgtgcctggccagatttttctttttattcttctttctttttttctt 7210
          |||  |||||||||||| ||| |||| | |  || |||| ||||||   |    ||||
Db   1595 CAGGCGTGAGCCACCGCGCCCGGCCTGGTCCTTTTTTTTTTTCTTCAGGCGGAGTCTT 1539


RESULT  11
```

```
PCT-US95-07201-10/c
; Sequence 10, Application PC/TUS9507201
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, Sofia
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION GENOMIC
;    TITLE OF INVENTION:  ORGANIZATION AND SEQUENCE OF THE PEDF GENE
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07201
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/367,841
;      FILING DATE:  30-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/257,963
;      FILING DATE:  07-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/952,796
;      FILING DATE:  24-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36434
;      REFERENCE/DOCKET NUMBER:  20264126PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  10:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7210 Base Pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Genomic DNA
;    ORIGINAL SOURCE:
;      ORGANISM:  Human
;    IMMEDIATE SOURCE:
;      LIBRARY:  DASH II
;    FEATURE:
;      NAME/KEY:  JT6A
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  7.0 kb Not 1-Not
;      OTHER INFORMATION:  fragment; Derived from human placental
;      OTHER INFORMATION:  genomic DNA; also referred to as JT106
PCT-US95-07201-10


  Query Match             2.2%;  Score 195.8;  DB 4;  Length 7210;
  Best Local Similarity   73.7%;  Pred. No. 1.7e-32;
  Matches  263;  Conservative    0;  Mismatches   92;  Indels    2;  Gaps    1;

Qy   6854 tttctctgttgtttggttcagggattttatttttaatttttatgagacagggtcttgctct 6913
          ||||| | |    |  ||       ||| | ||   |||| |||| ||| ||||||||||
Db   1893 TTTCTGTATCTGGTCCTTTTTTTTTATTTTTATTCTTTTTTTTGAGGCAGAGTCTTGCTCT 1834

Qy   6914 gtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctggg 6973
          ||| |||||||||||||||||||| |||||    | ||||||| ||||   ||||||||
```

```
Db   1833 GTCGCCCAGGCTGGAGTGCAGTGGCGTGATCTCGGTTCACTGCAACCTCTGCCTCCTGGG 1774

Qy   6974 ctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccg 7033
          |||||||||  ||||||||||||||| |  | |||||| |||||||| || ||||||| |
Db   1773 GTCAAGCAATTATCCTGCCTCAGCCTCCTGAGTAGCTGAGACTACAGGCGCCCGCCACTG 1714

Qy   7034 tgcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctg 7093
            ||  ||||| ||||   |||||  |||| |||||||| ||    | ||| | | ||||
Db   1713 CACCCAGCTAATTTTTGTATTTTTCAGTAGAGACGGGGTTTCACCATTTTGGCCAGGCTG 1654

Qy   7094 gtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattg 7153
          ||||  ||||  || ||||| |  |||||||  ||| ||||||||||||||||| ||||
Db   1653 GTCTTGAACTCCTGACCTCATG--ATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTA 1596

Qy   7154 cagatgtgagccaccgtgcctggccagatttttctttttattcttctttctttttctt 7210
          |||  |||||||||||| ||| |||| | |  || |||| |||||   |    ||||
Db   1595 CAGGCGTGAGCCACCGCGCCCGGCCTGGTCCTTTTTTTTTTCTTCAGGCGGAGTCTT 1539


RESULT  12
PCT-US95-07201-43/c
; Sequence 43, Application PC/TUS9507201
;  GENERAL INFORMATION:
;    APPLICANT:  Chader, Gerald J.; Becerra, Sofia
;    APPLICANT:  Patricia; Schwartz, Joan P.;
;    APPLICANT:  Taniwaki, Takayuki
;    TITLE OF INVENTION:  PIGMENT EPITHELIUM
;    TITLE OF INVENTION:  DERIVED FACTOR: CHARACTERIZATION GENOMIC
;    TITLE OF INVENTION:  ORGANIZATION AND SEQUENCE OF THE PEDF GENE
;    NUMBER OF SEQUENCES:  43
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Morgan & Finnegan, L.L.P.
;      STREET:  345 Park Avenue
;      CITY:  New York
;      STATE:  New York
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy Disk
;      COMPUTER:  IBM PC Compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/07201
;      FILING DATE:  06-JUN-1995
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/367,841
;      FILING DATE:  30-DEC-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/257,963
;      FILING DATE:  07-JUN-1994
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  07/952,796
;      FILING DATE:  24-SEP-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  DOROTHY R. AUTH
;      REGISTRATION NUMBER:  36434
;      REFERENCE/DOCKET NUMBER:  20264126PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212) 758-4800
;      TELEFAX:  (212) 751-6849
;  INFORMATION FOR SEQ ID NO:  43:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  22481 Base Pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Double
;      TOPOLOGY:  Unknown
;    MOLECULE TYPE:  Genomic DNA
;    FEATURE:
;      NAME/KEY:  Pl-147
;      LOCATION:
;      IDENTIFICATION METHOD:
;      OTHER INFORMATION:  full length genomic
;      OTHER INFORMATION:  sequence for PEDF plus flanking sequences.
PCT-US95-07201-43


  Query Match             2.2%;  Score 195.8;  DB 4;  Length 22481;
  Best Local Similarity   73.7%;  Pred. No. 2.9e-32;
  Matches  263;  Conservative    0;  Mismatches   92;  Indels    2;  Gaps    1;

Qy   6854 tttctctgttgtttggttcagggatttttatttttaattttatgagacagggtcttgctct 6913
          ||||| | |    |  ||       ||| | ||   |||| |||| ||| ||||||||||
Db   1885 TTTCTGTATCTGGTCCTTTTTTTTTATTTTTATTCTTTTTTTTGAGGCAGAGTCTTGCTCT 1826

Qy   6914 gtcccccaggctggagtgcagtggcatgatcatagctcactgcagcctcatactcctgggg 6973
          ||| |||||||||||||||||||||| |||||   | |||||||| ||||   ||||||||
Db   1825 GTCGCCCAGGCTGGAGTGCAGTGGCGTGATCTCGGTTCACTGCAACCTCTGCCTCCTGGG 1766

Qy   6974 ctcaagcaatcctcctgcctcagccttccaactagctgggactacaggtgcgcgccaccg 7033
          |||||||||  ||||||||||||||| |  | |||||| |||||||| || ||||||| |
Db   1765 GTCAAGCAATTATCCTGCCTCAGCCTCCTGAGTAGCTGAGACTACAGGCGCCCGCCACTG 1706

Qy   7034 tgcctggctaacttttcatttttttgtagggacggggtctcgttttgttgccaaagctg 7093
            ||  ||||| ||||   |||||  |||| |||||||| ||    | ||| | | ||||
Db   1705 CACCCAGCTAATTTTTGTATTTTTCAGTAGAGACGGGGTTTCACCATTTTGGCCAGGCTG 1646

Qy   7094 gtctcaaacttgtggcctcaagcaatccacctgccttggcctcccaaagtgctgagattg 7153
          ||||  ||||  || ||||| |  |||||||  ||| ||||||||||||||||| ||||
Db   1645 GTCTTGAACTCCTGACCTCATG--ATCCACCCACCTCGGCCTCCCAAAGTGCTGGGATTA 1588

Qy   7154 cagatgtgagccaccgtgcctggccagatttttctttttattcttctttctttttctt 7210
          |||  ||||||||||| ||| |||| | |  || |||| |||||   |    ||||
Db   1587 CAGGCGTGAGCCACCGCGCCCGGCCTGGTCCTTTTTTTTTTCTTCAGGCGGAGTCTT 1531


RESULT  13
US-08-451-777A-7/c
; Sequence 7, Application US/08451777A
; Patent No. 5789223
;  GENERAL INFORMATION:
;    APPLICANT:  Bergsma, Derk J.
;    APPLICANT:  Stambolian, Dwight
;    TITLE OF INVENTION:  Human Galactokinase Gene
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  709 Swedeland Road/UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/451,777A
;      FILING DATE:  26-MAY-1995
;      CLASSIFICATION:  436
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10825
;      FILING DATE:  23-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Eagle, Alissa M.
;      REGISTRATION NUMBER:  37,126
;      REFERENCE/DOCKET NUMBER:  P50268-1B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5364
;      TELEFAX:  610-270-5090
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7676 base pairs
```

```
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-451-777A-7


  Query Match             2.2%;  Score 195.2;  DB 1;  Length 7676;
  Best Local Similarity   80.5%;  Pred. No. 2.4e-32;
  Matches  243;  Conservative    0;  Mismatches   53;  Indels    6;  Gaps    1;

Qy   4312 tgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgggaggcagaggcag 4371
          ||  ||| | |||| |||||||||| |||| |||||||||||||||||| || |||||||
Db   6063 TGAATGGGTGGGCGCGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAAGCTGAGGCAG 6004

Qy   4372 gcagagcacttgaggtcaggaattcgagatcagcctggccaacatgaggaaaccccgtct 4431
          |||||   | ||||||||||  |||||| ||||||||||||||||  ||||||||||
Db   6003 GCAGATTGCCTGAGGTCAGGAGCTCGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCT 5944

Qy   4432 ctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgtaatcccagctactt 4491
          |||||||||||||||| | |||| ||||| |||||||| | ||||||||||||||||
Db   5943 CTACTAAAAATACAAAAGTTAGCCGGGCATCGTGGCAGGCGCCTGTAATCCCAGCTACTC 5884

Qy   4492 gggaggctgaggcaggagaat------cagaggggaggcggaggttgcagtgagccaaga 4545
          ||||||||||||||||||       |  ||||||||||||||||||||||   |||
Db   5883 AGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGTGGAGA 5824

Qy   4546 tcacgccactacaccccagcctaggtgacaaagcgagacttctcaaatattaacaataat 4605
          ||| |||||| ||| |||||| ||||||| || ||| || |      |  || || ||
Db   5823 TCATGCCACTGCACTTCAGCCTGGGTGACAGAGTGAGGCTCCCTCTCAAAAAAAAAAAG 5764

Qy   4606 aa 4607
          ||
Db   5763 AA 5762


RESULT  14
US-08-451-778A-7/c
; Sequence 7, Application US/08451778A
; Patent No. 5830649
;  GENERAL INFORMATION:
;    APPLICANT:  Bergsma, Derk J.
;    APPLICANT:  Stambolian, Dwight
;    TITLE OF INVENTION:  Human Galactokinase Gene
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  709 Swedeland Road/UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/451,778A
;      FILING DATE:  26-MAY-1995
;      CLASSIFICATION:  800
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US94/10825
;      FILING DATE:  23-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Eagle, Alissa M.
;      REGISTRATION NUMBER:  37,126
;      REFERENCE/DOCKET NUMBER:  P50268-1B
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  610-270-5364
;      TELEFAX:  610-270-5090
;  INFORMATION FOR SEQ ID NO:  7:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  7676 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-451-778A-7


  Query Match             2.2%;  Score 195.2;  DB 2;  Length 7676;
  Best Local Similarity   80.5%;  Pred. No. 2.4e-32;
  Matches  243;  Conservative    0;  Mismatches   53;  Indels    6;  Gaps    1;

Qy   4312 tgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgggaggcagaggcag 4371
          ||  ||| | |||| |||||||||| |||| |||||||||||||||||| || |||||||
Db   6063 TGAATGGGTGGGCGCGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAAGCTGAGGCAG 6004

Qy   4372 gcagagcacttgaggtcaggaattcgagatcagcctggccaacatgaggaaaccccgtct 4431
          |||||   | ||||||||||  |||||| ||||||||||||||||  ||||||||||
Db   6003 GCAGATTGCCTGAGGTCAGGAGCTCGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCT 5944

Qy   4432 ctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgtaatcccagctactt 4491
          |||||||||||||||| | |||| ||||| |||||||| | ||||||||||||||||
Db   5943 CTACTAAAAATACAAAAGTTAGCCGGGCATCGTGGCAGGCGCCTGTAATCCCAGCTACTC 5884

Qy   4492 gggaggctgaggcaggagaat------cagaggggaggcggaggttgcagtgagccaaga 4545
          ||||||||||||||||||       |  ||||||||||||||||||||||   |||
Db   5883 AGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGTGGAGA 5824

Qy   4546 tcacgccactacaccccagcctaggtgacaaagcgagacttctcaaatattaacaataat 4605
          ||| |||||| ||| |||||| ||||||| || ||| || |      |  || || ||
Db   5823 TCATGCCACTGCACTTCAGCCTGGGTGACAGAGTGAGGCTCCCTCTCAAAAAAAAAAAG 5764

Qy   4606 aa 4607
          ||
Db   5763 AA 5762


RESULT  15
US-08-998-208-7/c
; Sequence 7, Application US/08998208
; Patent No. 5880105
;  GENERAL INFORMATION:
;    APPLICANT:  Bergsma, Derk J.
;    APPLICANT:  Stambolian, Dwight
;    TITLE OF INVENTION:  Human Galactokinase Gene
;    NUMBER OF SEQUENCES:  33
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  SmithKline Beecham Corp./Corporate
;      ADDRESSEE:  Intellectual Property
;      STREET:  709 Swedeland Road/UW2220
;      CITY:  King of Prussia
;      STATE:  Pennsylvania
;      COUNTRY:  USA
;      ZIP:  19406-0939
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/998,208
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/451,777
;      FILING DATE:  26-MAY-1995
;      APPLICATION NUMBER:  PCT/US94/10825
;      FILING DATE:  23-SEP-1994
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Eagle, Alissa M.
;      REGISTRATION NUMBER:  37,126
;      REFERENCE/DOCKET NUMBER:  P50268-1B
```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5364
; TELEFAX: 610-270-5090
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7676 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-998-208-7

Query Match 2.2%; Score 195.2; DB 2; Length 7676;
Best Local Similarity 80.5%; Pred. No. 2.4e-32;
Matches 243; Conservative 0; Mismatches 53; Indels 6; Gaps 1;

```
Qy   4312 tgtgtggctaggcgtggtggctcacgcctgcaatcccagcactttgggaggcagaggcag 4371
          ||  ||| | |||| |||||||||| |||| |||||||||||||||||| || |||||||
Db   6063 TGAATGGGTGGGCGCGGTGGCTCACACCTGTAATCCCAGCACTTTGGGAAGCTGAGGCAG 6004

Qy   4372 gcagagcacttgaggtcaggaattcgagatcagcctggccaacatgaggaaaccccgtct 4431
          |||||   | ||||||||||| |||||| |||||||||||||||| ||||||||||||
Db   6003 GCAGATTGCCTGAGGTCAGGAGCTCGAGACCAGCCTGGCCAACATGGCGAAACCCCGTCT 5944

Qy   4432 ctactaaaaatacaaaaatcagccaggcatggtggcaggtgtctgtaatcccagctactt 4491
          |||||||||||||||| | |||| ||||| |||||||| | ||||||||||||||||
Db   5943 CTACTAAAAATACAAAAGTTAGCCGGGCATCGTGGCAGGCGCCTGTAATCCCAGCTACTC 5884

Qy   4492 gggaggctgaggcaggagaat------cagaggggaggcggaggttgcagtgagccaaga 4545
           ||||||||||||||||||||       |   ||||||||||||||||||||||   |||
Db   5883 AGGAGGCTGAGGCAGGAGAATTGCTTGAACCTGGGAGGCGGAGGTTGCAGTGAGTGGAGA 5824

Qy   4546 tcacgccactacaccccagcctaggtgacaaagcgagacttctcaaatattaacaataat 4605
          ||| |||||| ||| |||||| ||||||| || ||| || |      |  || || ||
Db   5823 TCATGCCACTGCACTTCAGCCTGGGTGACAGAGTGAGGCTCCCTCTCAAAAAAAAAAAAG 5764

Qy   4606 aa 4607
          ||
Db   5763 AA 5762
```

Search completed: April 6, 2001, 00:19:59
Job time: 17054 sec

Fri Apr 6 08:40:34 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:52:25 ; Search time 318.41 Seconds
(without alignments)
10674.903 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 480022 seqs, 187831343 residues

Total number of hits satisfying chosen parameters: 960044

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_36:*

RESULT 1
T23583
ID T23583 standard; cDNA to mRNA; 393 BP.
XX
AC T23583;
XX
DT 02-SEP-1996 (first entry)
XX
DE Human gene signature HUMGS05436.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN WO9514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
XX
PI Matsubara K, Okubo K;
XX
DR WPI; 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
PS Claim 1; Page 1404; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in T19001-T26837 and which is able to hybridise to part of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 393 BP; 73 A; 117 C; 121 G; 78 T; 4 other;

Query Match 4.0%; Score 365.8; DB 16; Length 393;
Best Local Similarity 97.2%; Pred. No. 2.1e-56;
Matches 381; Conservative 0; Mismatches 10; Indels 1; Gaps 1;

```
Qy   8454 gatcggactctgcccctggacctgggaacgactggactgtcacggggttccctcctagct 8513
          ||||||||||||||||||||||||||||||||||||||||||||||||||| ||||||||||
Db      1 gatcggactctgcccctggacctgggaacgactggactgtcacggggttacctcctagct 60

Qy   8514 ctcccagtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggact 8573
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     61 ctcccagtgaactcctgccaggcacacacagcccctatagcactgagctcacatgggact 120

Qy   8574 gggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggcccc 8633
          |||||||||||||||||||||||||||||||||||||||||||||||||||||||||| |||
Db    121 gggatatgggggcatctcttccccagagaggcactcagtgagcctcctgtgcctggnccc 180

Qy   8634 agtctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggagg 8693
          || ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    181 agcctgggccatctcttaggtgagacagttgcccgaaactaagccaggcctggctggagg 240

Qy   8694 agcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgct 8753
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    241 agcagcagcttggggagagggatttccctgcagacctcaagccatcatgcggtgggtgct 300

Qy   8754 gccatgacagaggctgcacccctgggccagcggggctgctcacccacctcttgtgcaagg 8813
          |||||||||||||||| ||| |||||||||||||| |||||| |||||| ||||||||||
Db    301 gccatgacagaggctggaccnctgggccagcggggttgctcaaccacctttgtgcaagg 360

Qy   8814 tggcctttgtgc-tgcgcctgcaggcagagct 8844
          |||||||||||| |||||||||||||||||  |
Db    361 tggcctttgtgcttgcgcctgcaggcagangt 392
```

FILE COPY

~~Fri Apr~~ 6 08:40:35 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 5, 2001, 19:21:50 ; Search time 4648.59 Seconds
(without alignments)
13639.351 Million cell updates/sec

Title: US-09-513-888-1
Perfect score: 9048
Sequence: 1 gcctttccaagaccctgccc.....tgccatttctcacgccctct 9048

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 7991742 seqs, 3503743858 residues

Total number of hits satisfying chosen parameters: 15983484

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*

RESULT 6
N21184/c
LOCUS N21184 619 bp mRNA EST 19-DEC-1995
DEFINITION yx41a10.s1 Soares melanocyte 2NbHM Homo sapiens cDNA clone IMAGE:264282 3', mRNA sequence.
ACCESSION N21184
VERSION N21184.1 GI:1126354
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 619)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
TITLE The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

High quality sequence stops: 342
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 2377 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 342.

FEATURES Location/Qualifiers
source 1. .619
/organism="Homo sapiens"
/db_xref="GDB:3873924"
/db_xref="taxon:9606"
/clone="IMAGE:264282"
/clone_lib="Soares melanocyte 2NbHM"
/sex="Male"
/tissue_type="melanocyte"
/lab_host="DH10B (ampicillin resistant)"
/note="Vector: pT7T3D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTGAAGTGGGAGCGGCCGCAGTTTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library constructed by Bento Soares and M.Fatima Bonaldo. RNA from normal foreskin melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."
BASE COUNT 165 a 163 c 160 g 125 t 6 others
ORIGIN

Query Match 5.6%; Score 502.4; DB 142; Length 619;
Best Local Similarity 93.5%; Pred. No. 1.9e-41;
Matches 576; Conservative 0; Mismatches 32; Indels 8; Gaps 5;

```
Qy   8427 agcagcactccgcatcacctcaccaaggatcggactctgcccctggacctgggaacgact 8486
          ||||||| || | || || |||||| |     |   |     ||| ||| ||||||||||||
Db    608 AGCAGCANTCGGAATAACTTCACCAGGATCGGATNT---GCCCCTGACNTGGGAACGACT 552

Qy   8487 ggactgtcacggggttccctcctagctctcccagtgaactcctgccaggcacacacagcc 8546
          ||||| |||| ||||||||||||||||| | ||  ||||||||||||||||||||||||
Db    551 GGACTTTCAC-GGGTTCCCTCCTAGCTTTTCCNATGAACTCCTGCCAGGCACACACAGCC 493

Qy   8547 cctatagcactgagctcacatgggactgggatatgggggcatctcttccccagagaggca 8606
          |||||||||||| ||||||||||||      ||||||||||||||||||||||||||||
Db    492 CCTATAGCACTGACTTCACATGGGACTGG--ATATGGGGCATCTCTTCCCCAGAGAGGCA 435

Qy   8607 ctcagtgagcctcctgtgcctggccccagtctgggccatctcttaggtgagacagttgcc 8666
          |||||||||| |||||||||| ||||||| ||||||||||||||||||||||||||||||
Db    434 CTCAGTGAGCNTCCTGTGCCT-GCCCCAGCCTGGGCCATCTCTTAGGTGAGACAGTTGCC 376

Qy   8667 cgaaactaagccaggcctggctggaggagcagcagcttggggagagggatttccctgcag 8726
          |||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Db    375 CGAAACTAAGCAAGGCCTGGCTGGAGGAGCAGCAGCTTGGGGAGAGGGATTTCCCTGCAG 316

Qy   8727 acctcaagccatcatgcggtgggtgctgccatgacagaggctgcacccctgggccagcgg 8786
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    315 ACCTCAAGCCATCATGCGGTGGGTGCTGCCATGACAGAGGCTGCACCCCTGGGCCAGCGG 256

Qy   8787 ggctgctcacccacctcttgtgcaaggtggcctttgtgctgcgcctgcaggcagagctgg 8846
          |||||||||||||||||||||||||||||||||  ||||||||||||||||||||||||
Db    255 GGCTGCTCACCCACCTCTTGTGCAAGGTGGCCTTTGT-NTGCGCCTGCAGGCAGAGCTGG 197

Qy   8847 agcccccagcagaggcaggctgggacggaccagcatctggaagatgtacatagttatttt 8906
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    196 AGCCCCCAGCAGAGGCAGGCTGGGACGGACCAGCATCTGGAAGATGTACATAGTTATTTT 137

Qy   8907 tctctttgtggtttcttgtttggtttggtttgcttttgacagcttcattttatttttgac 8966
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db    136 TCTCTTTGTGGTTTCTTGTTTGGTTTGGTTTGCTTTTGACAGCTTCATTTTATTTTTGAC 77

Qy   8967 gtcacttttttggccatgtaaactatttgtggcaattttatgttttttatttatgaataaag 9026
          ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db     76 GTCACTTTTTGGCCATGTAAACTATTTGTGGCAATTTTATGTTTTTATTTATGAATAAAG 17

Qy   9027 aatgccatttctcacg 9042
          ||||||||||||||||
Db     16 AATGCCATTTCTCACG 1
```

FILE COPY